

9140

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 2001-128126

(43)Date of publication of application : 11.05.2001

(51)Int.Cl.	H04N 5/92
	H03M 7/30
	H04N 5/765
	H04N 5/781
	H04N 5/84
	H04N 5/91
	H04N 5/93

(21)Application number : 2000- 278022	(71)Applicant : TOSHIBA CORP TIME WARNER ENTERTAINMENT CO LP
(22)Date of filing : 19.04.1996	(72)Inventor : YOGESHWAR JAY NG SHEAU-BAO ICHIGAWA TEIICHI UNNO HIROAKI MIMURA HIDENORI KITAMURA TETSUYA CHRISTOPHER J COOKSON THAGARD GREG B ROSEN ANDREW DRUSIN

(30)Priority

Priority number : 1995 502012 Priority date : 31.07.1995 Priority country : US

(54) METHOD AND SYSTEM FOR REPLACING SECTION OF ENCODED VIDEO
BIT STREAM

(57)Abstract:

PROBLEM TO BE SOLVED: To prevent degradation of a picture quality at the time of editing.

SOLUTION: A start point and an end point in an encoded bit stream are determined in order to remove a section of an encoded video. Start point and end point addresses of encoded data are calculated by summing up the number of bits consumed by each picturethe number of bits of a sequence header of each

picture the number of bits of each picture group (GOP) header and the total number of stuff bits. Processing to prevent decoding of unrelated signal components is executed in order to encode a video to be inserted. This processing is performed by judging the last picture before an editing point and using this last picture as a reference frame related to frames after the editing point. A short period of the first video encoded by the first quantization value is included in the end of the replaced video.

CLAIMS

[Claim(s)]

[Claim 1] A method characterized by comprising the following of adding video to a bit stream of encoded video.

A process of determining the starting point and a terminal point where said encoded video is corrected.

A process of obtaining unencoded video containing video in which video for inserting between said starting point and a terminal point and encoded video immediately after said terminal point are shown.

A process of assigning an additional bit to a bidirectional prediction frame which encodes a picture (GOP) of the 1st group of said video for insertion without using a reference frame of the outside of said starting point and a terminal point and does not contain the 2nd reference frame.

A process encoded to said terminal point of video in which GOP after said 1st GOP is inserted And a process of encoding said unencoded video by performing a process of encoding video in which encoded video immediately after said terminal point is shown a process of boiling encoded video which is created by said each process to encode in said encoded bit stream and arranging it.

[Claim 2] A method which is the method according to claim 1 and is characterized by including further a process of determining said additional number of bits using the number of bits used for a frame which has one-way prediction.

[Claim 3] A process of encoding video in which it is the method according to claim 1 and encoded video immediately after said terminal point is shown How including a process of encoding video in which said encoded video is shown using a quantized value equal to a quantized value used during the encoding of the beginning of said unencoded video.

[Claim 4] A method which is the method according to claim 1 and is characterized by said process to determine including a process determined by adding the number of bits corresponding to a video frame without referring to directory information of said encoded video.

[Claim 5] A method which is the method according to claim 1 and is characterized by including further a process of obtaining video using a camera and a process of encoding obtained video and generating a bit stream of encoded data.

[Claim 6] In a memory including a data structure for storing video of an encoded bit stream which was created by processing of claim 1 (1) The field which stored two or more menu cells which are the data structures stored in said memory and are one unit of (a) information (b) The field for start cells of a title menu the field for start cells of (c) audio selection menu (d) the field for start cells of a menu for choosing sub picture information used for providing different information from said video by which the bit map was carried out used and (e) when two or more programs exist in said memory When two or more angle screens exist in the field for start cells used in order to choose two or more programs and the (f) aforementioned memory A data structure including the field of a start cell for angle menus used in order to choose camera angle (2) It is a menu cell information table containing a table entry to each menu displayed A menu corresponding to the (a) aforementioned table entry respectively said said table entry in a title menu a program menu an audio menu and a sub picture menu. Or the field which shows whether it is an angle menu the field which shows the number of (b) selections (c) The field which shows a system clock reference part of a start part of a corresponding menu cell at least (d) A menu cell information table including the field which shows a start address of a corresponding menu cell and the field which shows two or more blocks of a menu cell which carries out (e) correspondence (3) A data structure for storing information displayed about each menu (4) A memory possessing a data structure for storing an audio and a data structure for storing video of an encoded bit stream which was created by processing of (5) claim 1.

[Claim 7] A memory wherein it is the memory according to claim 6 and said memory is an optical disc.

[Claim 8] A memory comprising including a data structure for storing video of an encoded bit stream which was created by processing of claim 1:

- (1) It is a data structure for storing file management information and is this data structure (a) A file control table for storing an address and size information of information in said memory
- (b) It is a sequence information table which stores information which shows turn that each cell which is one unit of information is reproduced This sequence information table including a table entry to each sequence each table entry (b1) The field which stores an end part of a head of a connection form sequence which is a part of perfect form type sequences which end after being reproduced and two or more sequences which are reproduced by turn the central part of a connection form sequence and a connection form sequence.
- (b2) The field which stores two or more cells within said corresponding sequence.
- (b3) The field which stores regeneration time of said corresponding sequence and the field which stores a sequence reproduced after said (b4) corresponding sequence.

[Claim 9] The memory according to claim 8 wherein said memory is an optical disc.

[Claim 10] A memory comprising including a data structure for storing video of an encoded bit stream which was created by processing of claim 1:

The field for being a data structure which this data structure includes structure for storing descriptive information including an audio and timing information and includes (1) descriptive information and storing "refer to the system clock of the (a) aforementioned descriptive information data storage."

(b) The field for storing time of onset of a corresponding video picture decoded without referring to other pictures.

(c) The field for storing a time stamp of audio information which has the reproduction time of onset in front of reproduction time of onset of said corresponding video picture.

(d) The field for it being shown whether the 1st bit is located in whether said audio signal is located before said descriptive information data structure and the back and storing an address of said audio information (e) A data structure including the field which shows an address of other descriptive information data structures generated before said descriptive information data structure and in the back (2) a data structure for storing an audio and (3) -- a data structure which stores video of an encoded bit stream which was created by processing of said claim 1.

[Claim 11] A memory wherein it is the memory according to claim 10 and said memory is an optical disc.

[Claim 12] A memory comprising including a data structure for storing video of an encoded bit stream which was created by processing of claim 1:

The field for this data structure being a data structure stored in said memory since

(1) sub-picture unit is stored including structure for storing an audio and sub picture information and storing (a) sub picture unit header.

(b) A subfield which are the field for storing a picture by which the map was carried out and the field for storing a display control table containing a table entry of (c) plurality and stores display start time of said (c1) picture by which the bit map was carried out.

(c2) A subfield which stores an address of the following table entry And (c3) a data structure including a subfield which stores at least one display control command chosen from a group of commands including a color of a pixel produced when display area a pixel color pixel contrast and a video frame are displayed and contrast change.

(2) A data structure for storing video of encoding created by a data structure for storing an audio and processing of (3) claim 1 now a ** bit stream.

[Claim 13] A memory which is the memory according to claim 12 and is characterized by said sub picture unit data structure possessing two or more packs including information within said sub picture unit data structure.

[Claim 14] A memory wherein it is the memory according to claim 12 and said memory is an optical disc.

[Claim 15] it is a system which adds video to a bit stream of encoded video -- (1) -- with a means to determine the starting point by which said encoded video is corrected and a terminal point. (2) A means to obtain video in which encoded video

immediately after unencoded video containing video for inserting between said starting point and a terminal point and said terminal point is shown(3) Are said unencoded video a means to encode and this encode means(a) Encode a picture (GOP) of the 1st group of video for insertionwithout using a reference frame of the outside of said starting point and a terminal pointA means to assign an additional bit to a frame predicted by both directions which do not contain the 2nd reference frame(b) A means to encode two or more GOP(s) after said 1st GOP to a terminal point of video in which it is insertedand (c) -- a means to encode video in which encoded video immediately after said terminal point is shown is included --
- (4) -- a system possessing a means to arrange encoded video which was created by said each process to encode in said encoded bit stream.

DETAILED DESCRIPTION

[Detailed Description of the Invention]

[0001]

[Field of the Invention]Firstrelated application is described.

[0002]This application relates to the United States Patent specification sequence numbers 08/438014 entitled "the method and system" for exchanging the section of the encoded video bit stream which were submitted on May 81995. This application againThe United States Patent specification sequence numbers 08/467991 entitled "the audio image encoding system with which the number of audio encoders is reduced"the specification sequence numbers 08/466 entitled "the method for changing the quality of the already encoded video frame manuallyand a system"391 and the United States Patent specification sequence numbers 08/469370 entitled "a method for a user to change the quality of the already encoded video frame manuallyand a system"by changing a quantization level, Bit assignment of a video encoder. The rate controlling expression digital video editing method to control. It relates to the United States Patent specification sequence numbers 08/466,766 entitled and a systemand the United States Patent specification sequence numbers 08/473783 entitled "the video encoding method and system which are encoded using a rate quantization model"Both these patent specifications are submitted on June 61995and are taken in by this specification as reference.

[0003]The encoded bit stream section this invention about the encoding system transposed to a new sectionIt is related with the video encode system replaced so that the video sections encoded especially may decrease in number a decoding accessory constituent (decoding artifacts).

[0004]

[Description of the Prior Art]Before being able to use the compressing methodan audiovideoand a film only cut off an audiovideoand film[a sectioni.e.a part of]and edit was performed by being new or transposing the section to a substitute section if needed. Howeverwhen the compressing method which encodes data came to be usedsimple "cut and attachment" edit became impossible as

information used for constituting a video frame. Therefore cutting off the encoded video sections simply may degrade the video sections which are not changed and when new video is simply inserted in the already encoded bit stream it may contain a decoding accessory constituent.

[0005] In order to solve this problem in MPEG video encoding there is the method of carrying out label attachment as a group (GOP: group of pictures) of the picture which closed the new video sections replaced in former video. In MPEG encoding the group of the closed picture shows that the prediction used into B type frame just behind the I frame by which the 1st was coded uses only backward prediction. Or MPEG enables use of the flag called "broken_link." Since the 1st B frame just behind the I frame by which the 1st was coded cannot use the reference frame used for prediction when set up this flag shows that it may not be decoded correctly. That decoder uses this flag in order to prevent the display of the frame which cannot be decoded correctly.

[0006]

[Problem(s) to be Solved by the Invention] The above-mentioned solution degrades the encoded video and may provide the low video of quality.

[0007] Then the already encoded data section is a system transposed to other encoded data sections and an object of this invention is to provide the new method and device of the decoding accessory constituent which is visible to eyes which are not [that it is few or].

[0008] The purpose of this invention so that the encoded video sections which should be replaced may not include an error i.e. a video decoding accessory constituent in the already encoded video. It is providing the video encode system replaced with other encoded video sections.

[0009] An object of this invention is to provide the memory which has a data structure which stores information by a new and useful method.

[0010]

[Means for Solving the Problem] The above and other purposes are attained by a new method and a system which replace a section of an encoded bit stream. In order to remove a section of encoded video it is necessary to judge the starting point and a terminal point of video which were encoded within a bit stream encoded in digital one. This is performed by adding the number of bits used for each already encoded frame in order to judge bit offset over the target screen. A directory can be used in order to maintain information which shows an exact position or time of a specific frame in a bit stream encoded as other methods.

[0011] The 1st section of one GOP which is 1/2 of original video for example the 2nd section. It is located in the starting point of the new video encoded and it is used in order to generate p frames of the last of GOP in front of a cutting point required to encode and decode the first two B frames of the next GOP correctly. Data encoded from this short initial section is not replaced within already encoded data and it is used only in order to generate a reference frame (the last p frames) referred to by two B frames of the beginning after a start point.

[0012] A short section of said original video is located in a terminal point of

video and is encoded using a quantized value used when encoded first. Since a frame of encoded video in GOP returns to an already encoded frame and is referred to when a frame of the first encoding returns to the new frame of video in front of said terminal point and is referred to sudden fault about decoding may occur. Therefore it is desirable for original video to return to a frame equivalent on an original frame and to refer to it. This is not essential although one half of original videos is again encoded at the starting point of an edit segment.

[0013] As other method substitution video is encoded using a group mode of a picture which common knowledge closed so that a frame in substitution video may not be dependent on a frame of the outside of substitution video. As a spare bit is assigned to a bidirectional frame as for the new feature of this method and the number of bits of the B frame shows a rough value of the number of bits used for P frames this bidirectional frame is having only unidirectional prediction for closed GOP mode. In this method closed GOP mode does not contain a big decoding accessory constituent in the B frame but the B frame has the same quality as about P frames.

[0014] Encoded video which was stored in a digital storing medium also constitutes a part of this invention. This invention contains a memory which stores a data structure new and useful again. This memory is a suitable optical disc.

[0015]

[Embodiment of the Invention] Hereafter this embodiment of the invention is described.

[0016] This invention is a digital video encoding system provided with two or more components and functions. This specification is composed by the paragraph shown below in order to identify various components of a system still more easily.

I. It is [The log file for memorizing the information on the encoded video] II. system component III. first. General-purpose system action IV. System substitution measure V. VI. which reduces the number of audio encoders while maintaining the optimal encoding efficiency And log interpreter utility VII. for accessing a log file and the encoded video Change A. of the quality of the encoded video Change B. of the quality in a period Change VIII. of the quality in the field of a frame. Re-encoding A. of the video after quality change Bit quota control B. when quality is changed within an independent frame 1. re-encoded while avoiding decoding on an edit point When it encodes first. 2. which restores an encoder to the state where it was used having been alike Re-encoding IX. in the closing mode of the group of a picture The rate quantization modeling X. data structure used for encoding is used An audio video and the system indicated at the beginning of I [which combines the digital signal of a sub picture]. are related with the digital audio image encoding system known also as a digital compression system. In the idea indicated by this specification are usable with all kinds of encoding scheme Especially Formally as MPEG-1 (ISO/IEC 11172-video) which is the past of ISO/IEC 13818-2 ISO/IEC [which specifies MPEG-2 system information] 13818-1 and MPEG-2. It is applicable to MPEG-2 (MPEG: Moving Picture Experts Group) known. In this specification MPEG-1 and MPEG-2 are collectively called "MPEG."

Although the MPEG standard specifies only the syntax of the video stream compressed or encoded how compression is performed does not specify it. Since the data stream by which MPEG encoding was carried out is a compressed data stream an encoding process can be called a compression process and the reverse should also be noticed about the possible thing. A system is related also with MPEG-2 standard defined as the encoding 13818-3 of the audio which can be performed using Dolby (registered trademark) AC-3 i.e. ISO/IEC. Each of said standard shall make some of these specifications by referring to them.

[0017] MPEG video compression is the predictive compression of time causal relationship. An MPEG encoder is compared with the predicted picture in which the encoder created a new video picture based on the already encoded picture if a new video picture is received. Since the term of "causal relationship (causal)" prediction coding is used since prediction is stood based on the information transmitted before and prediction is performed by a time picture unit the term of unfortunate prediction coding "time (temporal)" is used. Movement of the field within a video picture needs to be predicted for a prediction procedure. Therefore time prediction may be called motion compensation prediction.

[0018] The result of comparison of the old picture and a new picture expresses the new information stored in a next video picture. And the new information called "remains information (residual information)" will receive a series of information-compression processes. First primary mathematics conversion called a discrete cosine transform (DCT) is performed. Pixel remains information is changed into the sequence of a coefficient by DCT operation. In MPEG compression DCT is performed with the block which comprises 8x8 pixels. Similarly a DCT coefficient is also put in in the number of 8x8 arrangement. Then these coefficients are individually quantized to the grade to which accuracy was set based on quantization step size i.e. q-level (that is rounded off). A quantization process usually produces many coefficients to which value zero were set. Otherwise the coding structure called quantization noise makes it this quantization process produced.

[0019] The quantized coefficient is coded after that using the combination of the variable length coding of two numbers of groups with which run length coding by the coefficient to which the zero value was set and the run length value which each produces as a result and coefficient values other than zero are expressed respectively. The code of the group of these run length nonzero values forms the compressed digital data stream corresponding to a video picture with other overhead information. It is required to change the grade of quantization so that compression of a suitable quantity may be attained for physical restriction like an output data transfer rate or a storage capacity. This is called buffer feedback.

[0020] The decoder based on MPEG reads an encoding digital data stream and performs an operation contrary to an encoding process.

[0021] Generally since it is less than the bit total for which the bit total in a digital data stream is needed since a video picture is expressed on parenchyma compression is attained. However it is important to recognize that

compression which is illustrated by MPEG video coding never escapes distortion and artefact (artifacts). As mentioned above the main source of compression artefact is quantization processes.

[0022] The indicated system aims at adjusting the grade of quantization dynamically between a compression process or an encoding process. Especially adjustment is performed according to the directions issued by human being's operator. Adjustment may be applied to the whole video sequence if it may apply to the portion of a bigger video sequence.

II. a system component -- here reference of the drawing in which the drawing 1 A is shown when a similar reference number shows the same portion or a corresponding portion through two or more figures and specifies further shows the block diagram of the encoder system architecture. Discontinuation and in order to attain encoding and edit errorless it should annotate to it being one complicated process of needing the device with which many which operate synchronizing with a video encoding process differed. However audio encoding and video encoding can be separately executed as long as a suitable time code is used. Although four workstations are drawn on drawing 1 A it is usually only one keyboard and workstation that the user or operator of an encoding system has a dialog. A workstation possesses a processor RAM ROM a network interface a hard disk a keyboard a monitor and a conventional general purpose computer component like the other conventional component.

[0023] The workstation 30 functions in the example as a system control station which is a Sun SPARC20 Unix workstation. In the workstation 30 not only a conventional workstation input device like the keyboard 32 and the POINTIMBU device 33 but the graphic-console display monitor 31 which are a mouse and the other pointing device is provided. A user will mainly operate the workstation 30 visually via the graphic user interface (GUI) displayed on the console display 31.

[0024] It is connected to another workstations 40 10 and 20 explained below and the workstation 30 controls them. The workstation 30 is connected also to the audio interfacing unit 72 and the digital video capture memory storage 60 which are explained below too. The workstation 30 is connected to the workstation 40 via the serial interface defined by American Electronics Association (EIA) standard RS-232. Similarly connection between the workstation 30 and the audio interfacing unit 72 is also depended on an in-series RS-232 standard. The connection between the workstation 30 and the digital video capture memory storage 60 passes S-bus interface of the industry standard.

[0025] the frame jogger 41 -- an operator -- VTR 51 or 52 -- it is connected to the workstation 40 in order to be able to position one of videos in the part which must perform edit. In order to control the position of video the time code which an operator inputs can be used for a system. However the jog control device is raising the degree of user FRIENDLY of the system by enabling use of a jog knob in order to position video. The workstation 30 controls whether it is that the video displayed by the video display monitor 61 is digital capture memory storage 60 or recording VTR 62 or the video from the both.

[0026]The video encoding device 50 is a digital video encoder which performs compression according to MPEG-1 or MPEG-2 or the other video compression standard. MPEG-1 encoder marketed is Sony (registered trademark) RTE-3000. MPEG-2 encoder can be built according to the instruction explained in the Japanese patent specification 6-326435 for which it applied on December 27, 1994 taken in by this specification by reference. The workstation 10 controls the video encoding device 50 directly via the industry standard S-Bus and the workstation 10A command is received from the workstation 30 via a standard networking device like an Ethernet (registered trademark) network. The workstation 10 supports real-time execution of the video encoding control program of the video encoding device 50 between encoding processes. in an example -- a works -- the station 10 is a Sun SPARC20 Unix workstation.

[0027]The workstation 10 starts a compression process in response to the command from the central workstation 30. The workstation 10 controls dynamically the grade of the quantization applied to the specific field of a video signal between compression processes. The specific field to which quantization is changed explains below how a spatial field or its both and quantization are controlled strictly.

[0028]The input video to the video encoding device 50 is from the tape in playback VTR 51. Reproduction VTR 51 are a digital video standard for occupations and ITU-R 601 (called the CCIR 601 video standard before) video tape player. This standard is applied not only to a PAL system but to an NTSC system. It comments on the electromagnetism information showing the various pictures and scenes which the capture was carried out with the video camera or the film camera or were created by computer being memorized to the tape in playback VTR 51. The output encoded from the video encoding device 50 is transmitted to the workstation 20 in order to memorize with the one or more hard disks 21.

[0029]In order to display the video data encoded and outputted by the video encoding device 50 a digital video decoder like MPEG-2 is needed. However at the creation time of this specification as for the realized systems since it did not obtain easily between the initial development phases of a system such a decoder does not mount MPEG-2 separate decoder. Therefore the video encoding device 50 not only outputs MPEG-2 video stream to the workstation 20 but the decrypted equivalent video image is outputted to the encoding data of MPEG-2 from the video encoding device 50 recording VTR 52. It is connected to the video encoding device 50 by the digital video signal connector of the industry standard known as D1 interface and recording VTR 52 is controlled by the workstation 40. Both playback VTR 51 and recording VTR 52 are appropriately realized using the video tape recorder for Ampex DCT occupations.

[0030]An audio is encoded and compressed into the format based on MPEG-2 which is described by 13818 to ISO/IEC 3 standard for example instead of the Dolby AC-3 format. The audio sources of an encoding system are four digital audio tape (DAT) players and 71a, 71b, 71c and 71d. Sony is marketing the DAT player of studio quality. DAT player 71a-71d via industry standard protocol RS-

422it is connected by the workstation 40 and controlled by it. The audio signal outputted by DAT player 71a-71d is inputted into the audio encoders 70a70b70cand 70drespectively. It realizes using Dolby AC-3 model DP 525 commercial encoderand these audio encoders are controlled by the audio interfacing unit 72. Usuallyaudio encoding is started and ended with the specified time code. The audio interfacing unit 72 is on RS-232 connection in the meantimeand receives a command like an encoding start time code and the end time code of encoding from the workstation 30. The audio interfacing unit 72 multiplexes various audio information compressed [digital] to independent Digital Stream transmitted to the workstation 20. The workstation 20 carries out demultiplexing of the input digital audio stream to two or more files corresponding to the compressed digital data outputted by the audio encoding device 70. The video encoding device 50 provides the audio interfacing unit 72 with the synchronized signal based on the video synchronizing signals. The audio bit stream can synchronize now with the video bit stream outputted by the video encoding device 50 with a synchronized signal.

[0031]The workstation 30 is provided with the function it is directed to the audio interfacing unit 72 that provides the duplicate of audio information from audio encoding device 70a-70d various either to the audio decoder 73. The audio decoder 73 decrypts a compressed audio so that a user can hear an audio via the loudspeaker 75. Human being's operator opts for the selection about which audio encoding device is monitored through the graphic user interface on the console display 31 with either the keyboard 32 or the pointing device 33. The composition of the audio interfacing unit 72 is the hardware provided with the function to perform the above-mentioned function. This hardware is the conventional circuit element arranged in order to attain one or more the programmed microprocessors or the functions which were indicatedor its both.

[0032]The workstation 40 is a studio equipment control stationand in order to control the playback videotape recorder 51 and digital audio tape player 71a-71d both with the recording videotape recorder 52it publishes a suitable command. The workstation 40 is ordered to start the capture of video at the suitable time for the digital video capture memory storage 60. The connection between the workstation 40various videotape decksand an audio tape deck is a RS-422 protocol of an industry standard. If this protocol is usedeach tape deck can tell the workstation 40 about that present tape location in a standard time code format. The workstation 40 synchronizes various tape decks appropriately so that such information is usedit may be rightand may be played and the information on video and an audio may be recorded. In this examplethe workstation 40It is a standard IBM compatible personal computer which performs Ensemble Pro which is the software program which EditingTechnologies Corp.Ensemble Pro in DOS and Californiaand MOAPAKU is marketing. It comments on having added a minor improvement to the Ensemble Pro program so that the workstation 30 can communicate with Ensemble Pro currently performed by workstation 40 via a RS-232 communication port. It can either also market [this improvement] via

Editing Technologies and Corp. or perform not conducting an excessive experiment by a person skilled in the art.

[0033] The digital video capture memory storage 60 performs two functions shown below within an encoding system. For the 1st this device uses that capture memory storage function in order to perform "front and back" comparison with convenient manual edit applied to the video signal and it provides [2nd] a video display with a graphic overlay function. This graphic overlay function is used between space manual edit phases and enables it to define the field where human being's operator has the concern from which the number of the bits assigned as a result of [its] quantization changes for example.

[0034] In this invention digital video capture memory storage is realized as audio product model VS6000 of the DRABM base where Viewgraphics in California and Mountain View is marketed. The digital video inputted into this device is based on the industry standard D1 provided with ITU-R-601 (old CCIR 601) image resolution. The output of the digital video capture memory storage 60 is connected to a video display via three separate analog signals showing the red of a video signal and a green and blue component. Since graphic overlay information is processed by the digital video capture memory storage 60 before creation of the last analog signal outputted by the digital video capture memory storage 60 By the video display monitor 61 human being's operator can see with other video datas.

[0035] The workstation 30 which is operating in the Unix environment carries out the map of the graphic window after the video display monitor 61 which uses the digital video capture memory storage 60. By this human being's operator can perform now graphic (it is (like rectangular drawing)) operation in respect of graphic overlay of the display shown in the monitor 61. The device 60 realizes an overlay function. Before overlay changes the information in a memory into the analog video signal which drives a monitor it is performed by video memory. VS6000 which a system uses has 8-bit memory in the upper part of video memory. In respect of this 8-bit the computer can draw all color graphics and texts. One of the "colors" in this 8-bit side is a transparent color. The arbitrary pixels on this overlay side that is a transparent color take the value of the video memory in the bottom of it. Generally most overlay sides become a transparent color and graphics (a line a text etc.) use colors other than transparence. Therefore the display memory which consists of both video memory and an overlay graphic memory is formed. In order to display by the monitor 61 it is a display memory ultimately changed into an analog signal. human being's operator was usually connected with the keyboard 32 -- pointing device 33 W use of is done and such graphic operation and operation of a graphic object created by doing in this way are performed.

[0036] The workstation 20 receives a digital data stream not only from the audio interfacing unit 72 but from the video encoding device 50. The workstation 20 is connected to the workstation 30 by Ethernet connection and the workstation 20 is connected also to the video encoding device 50 and the audio interfacing unit 72 via S-bus interface of an industry standard. The received digital data stream is memorized in one or more hard disks as a separate data file. A separate Direct-

Memory-Access (DMA) card is used when transmitting the Digital Stream data by which the capture was carried out to the hard disk 21. When it is a long video sequence which a lot of digital data produces as a result the hard disk 21 can be realized using a hard disk like SPARCstorage Array X655 A/G5 made from Sun Microsystems marketed. In this example the workstation 20 is SPARKserver 1000 by which the product made from Sun Microsystems is marketed.

[0037]The workstation 20 also formats the data file corresponding to data compressed [video] and audio compression finishing data into the suitable independent format file for the transfer or transmission to a decoding device. It will be based on the disc format to which it is desirable that it is a disk base as for a final medium and a format relates in that case. Usually the formatted file is transported to the mechanism which transmission to one or more last transmission media generates by a certain middle digital memory measure like a digital tape. The digital memory storage 22 currently illustrated is used for such a purpose. In this example the digital memory storage 22 possesses 8 mm Exabyte tape drive marketed. The interface of the digital memory storage 22 is a peripheral equipment interface (SCSI) for small computers of an industry standard.

[0038]It is known that two or more SCSI devices are connectable with the same SCSI bus. Therefore the digital memory storage 22 can possess other SCSI devices like the digital linear tape (DLT) drive marketed a magneto-optics (MO) disk drive or a floppy (registered trademark) disk drive. These composition corresponds easily and is useful to provide the flexible input device for ancillary data which may be needed between the format stages before a format stage. For example in most cases in manufacture of a movie caption data can be used within the medium which needs one of the aforementioned SCSI devices. In such a case ancillary data is read by workstation 20 by the software program under execution and it is processed so that it may be stored in the file into which ancillary data was formatted. Ancillary data may store the title data which is a U.S. closing caption format. The data of hope like a program code or PostScript data can be put into ancillary data.

[0039]Drawing 1 B corresponds to the system of drawing 1 A and the audio / image encoding system 98 which functions as the system of drawing 1 A similarly show it with the block diagram. In drawing 1 B the input video of encoding is TEMUHE is supplied from the camera 80 and an audio is supplied from the microphone 82. Video information and the motion-picture film which probably memorizes audio information are transmitted to the encoding system 98 via the transmission device 88 possessing the photodetector for carrying out the capture of the picture from the light and film for illuminating a film. Before the information from the input devices 80 and 82 is encoded by the encoding system [like] 98 by the audio tape recorder or a videotape recorder it is memorized in an electromagnetism format. An audiotape and videotape are encoded by the system of drawing 1 A.

[0040]The final digital bit stream made by an encoding system It may be sent to the optical disc mastering device 90 which carries out press working of sheet

metal of the optical disc and the optical disk writing device 84 written in an optical disc may be transmitted to a television receiver or a set top box decoder via the television transmitter 86. Press working of sheet metal of an optical disc and creation are performed by the method marketed by the known which uses a master disc for example. The encoded information can be used also by a video-on-demand system. The encoded bit stream is ultimately decrypted by the decryption process corresponding to an encoding process and audio video information is expressed to a user as television or a monitor. A decoding device possesses the digital video disk player which decrypts the disk encoded in order to display by television or monitor.

[0041] The electrical signal with which an encoding system expresses an audio and a picture is inputted and an electrical signal is changed into a new format and ultimately it is clear from drawing 1 B that a signal is decrypted in order to re-create the information on the encoded origin.

III. System action drawing 2 is a flow chart explaining general operation of the system shown in drawing 1 with a block diagram. A system goes through two or more "states" between the motion cycle so that drawing 2 may show. A deep solid line shows a standard operation course and a solid line draws the end course of discontinuation and a dashed line shows a bypass route.

[0042] The first operation by a system is generated in the state of [100] a setup. Before the compression state 102 is inputted by this step an input parameter can be specified by it. In the compression state 102 the video image memorized by the tape in playback VTR 51 is encoded by the format by which digital compression was carried out with the video encoding device 50.

[0043] The user can edit a compression process manually after the compression state 102 based on the image quality of the data compressed in the past of the edited state 104. By this the user can improve or reduce image quality in the inside of a field of one frame of video or a period. The term of "edit" should be cautious of not meaning adding the scene itself to a picture stream or deleting. In this specification the term of "edit" means quantization and changing the quality of video continuously. After a picture is edited by the edited state 104 in order to encode data according to a new user specification parameter it is required to perform compression operations again at the compression state 102.

[0044] After a user checks the form that video was edited all the information including an audio like a title video and the other information is collectively combined with the format of hope by the format state 106. In the completion state 108 all temporary files are eliminated and a process is ended according to the exit status 110. Furthermore it is related with each state where it is explained to drawing 2 here explains detailed information with reference to drawing 6 from drawing 3.

[0045] Drawing 3 shows the step performed in the state of [100] a setup. Step 120 started the setup state and Step 122 initialized the system. In this step execution of the system control software stored in the workstation 30 begins. It is displayed on the console display 31 and the other program like the

graphic user interface (GUI) program which takes charge of the user interface into which a parameter is made to input via the keyboard 32 and the pointing device 33 is started. Between Step 122 the system controller software of the workstation 30 refers to other devices and traces the system state containing the available device of a system.

[0046] Information is expressed to a user as Step 124 through GUI. Existence of the parameter used for the information about the message of a welcome and a system and an encoding process and all the files encoded in the past is displayed. In the displayed system information the device connected to an audio encoder VTRan audio tape deck and a system including the composition is provided. When the already encoded file is stored in the system the information which describes the audio encoding parameter and video encoding parameter which were used for information creating a file is displayed on a user.

[0047] Encoding which uses a standard system parameter or a user specification system parameter in Step 126 The information from a user like the command for starting the check can accept the parameter to a system parameter changed or chosen is inputted.

[0048] And in Step 128 in order to prepare encode operation and decryption operation input data is processed by transmitting the data inputted into system controller software from the graphic user interface. A setup state is ended by 130.

[0049] A block diagram is shown for the process performed by the compression state 102 in drawing 4. According to the standard operation flow the user already specified the video and the accompanying audio sequence which are compressed in the state of a setup. And a user directs to perform an automatic procedure i.e. an automatic compression procedure to a system. A user is a period or cannot control a dynamic quantization process by video compression between automatic compression procedures in one frame of video. This is meant as initial compression execution with which a coding decision is made based on the objective standard calculated by the video encoding device 50. When automatic compression has already been performed a user is one of the spatial (on a frame or video) or time targets (with video sequence) and can adjust quantization by edit compressed mode manually.

[0050] In a compression state after starting at Step 140 it prepares for compression operations at Step 142. The parameter of a compression process like the accompanying audio compressed downloads from the workstation 30 to the video encoding device 50 via the workstation 10 with the time code and compression parameter of working [this] and video.

[0051] And the kind of compression of hope is determined in Step 144. When compressing information for the first time a compression process is automatically performed without a user's intervention. An automatic compression process is prepared in Step 146. During this preparation the workstation 30 determines which [of digital audio tape player 71a-71d and the related audio encoding device 70a-70d] must be used. The workstation 30 is ordered to prepare for the video encoding device 50 and the bit stream capture from an audio encoder 70a-70d

output which passed the audio interface unit 72 to the workstation 20. A tape is carried forward to a suitable starting position playback VTR 51 recording VTR 52 and DAT player 71a-71d. And the workstation 40 sends a signal so that movement at Step 148 may be started to the playback videotape recorder 51 the recording tape recorder 52 and DAT player 71a-71d. Then according to the format of hope like MPEG-1 which uses the video encoding device 50 or MPEG-2 format a video data is encoded at Step 150. Audio information is also encoded according to the format of hope like the Dolby AC-3 format. In order to assist a manual edit process later it is necessary to create a log file. A log file shows the parameter of the encoded video data which is needed in order to perform edit after a video data. An encoding process stops at Step 160 which reaches a suitable time code. If this time code may be the last of the information which should be encoded it may be a point specified by the user who is not the last of a video sequence.

[0052] So that an automatic compression process may already be performed and a user may explain drawing 5 below at Step 144 if it is judged that he wishes to change the image quality of one or more frames of video manually using the parameter specified between edited states it will be judged at Step 144 whether it is necessary to perform edit-compression operations. In Step 152 preparations of edit compression operations including the signal of reproduction VTR 51 to the start points of the hope which is a point which manual edit starts are made. Since it is not necessary to change it once audio information is encoded by automatic compression operations the disable of the audio equipment is carried out. And playback VTR 51 and recording VTR 52 move at Step 154.

[0053] In Step 156 encoding of the video data according to the parameter specified by the user by the edited state explained to drawing 5 begins after that. Edit compression is ended at Step 160 which reached the end time code. In Step 162 the message of the purport that compression and encoding were completed is displayed on a user and a process is ended at Step 164.

[0054] Drawing 5 explains the process performed by the edited state 104. As mentioned above the edit said here is not edit of the conventional video to which a scene is deleted shortened or moved within a video sequence. The edited state in this specification refers to changing the quality of some videos or the time sequence of video by setting up quantization of video manually.

[0055] After starting an edited state at Step 170 in Step 172 it is asked whether wish manual edit of the video by a user changing bit assignment of the encoded video. When a user does not desire change of bit assignment a process is ended at Step 188. When a user desires change of bit assignment a user defines the video segment of an editing object by Step 174. This is performed by choosing the period of the video which must change image quality. And it is judged at Step 176 whether a user wishes spatial edit or time edit. Time edit changes assignment of a bit in a period to spatial edit changing quantization or assignment of a bit in the frame of video. A user's selection of spatial edit will advance a flow to Step 178 which inputs the field in the frame edited. The relative grade of change applied is also inputted. In the case of this invention the integer graduation of -5 to 5 [+]

containing zero is used in order to show the relative amount of change. A user marks the field of video using the keyboard 32 the pointing device 33 or its both and specifies one of the integers which contain -5 and +5 from -5 to +5. The user can also show that the quantization level before set as a certain field must not be changed. For example users decrease in number bit assignment of the other field when wishing for the increase in bit assignment of a certain fixed field. If a field with a user is set up as change called also "with protection (protected)" is impossible the bit which is needed for an improvement of an image quality field will not be extracted from a field with protection.

[0056] If it is judged that a user wishes time edit at Step 176 a flow will progress to Step 180 which inputs how a period must be edited. In time edit as well as spatial edit a user specifies the integral value between -5 which shows the degree of relativity of change applied to a specific sequence with video and +5. This change becomes effective in the whole selected period.

[0057] After either spatial edit or time edit is performed at Steps 178 and 180 at Step 182 according to the parameter inputted manually a video scene is reencoded for example is recompressed into MPEG-2 format. Encoding-already data is displayed next to the data edited newly so that a user can compare in detail how image quality was changed.

[0058] If an operator shows the segment of an editing object the video from VTR 52 is transmitted to the memory storage 60 and even if there are many available total memories a half will be occupied with the memory storage 60. The segment is equivalent to a "front" segment. Between edit-compression processes by workstation 20 the capture of the bit stream is carried out and it is memorized as a separate bit stream file from a master video bit stream file. A master video bit stream file is a compressed bit stream of the whole movie. The video edited - compressed is recorded by VTR 52. Between encoding processes an operator will display a picture at the same time as it is recorded by VTR 52. When you wish the display of the picture after an operator encodes VTR can be used as reproduction machinery. When an operator thinks that he would like to perform comparison of "a front and the back" the segment to which VTR 52 corresponds is transmitted to the memory storage 60 and even the remaining half of the total memory of the device 60 is consumed.

[0059] An operator determines after this encoding whether it is necessary to save the changed video. When it is decided that an operator disregards a "next" segment a master bit stream is left behind without being changed and the bit stream edit-compressed is deleted. In order for VTR 52 to enable it to reflect a master bit stream another automatic-compression is performed by an edited segment by only the purpose of recording a "front" segment. In order to transmit a "front" segment to VTR 52 and to return the device 60 the alternative plan of using a device is also possible. When it is decided that an operator memorizes a "next" segment it is necessary to update a master bit stream file and to put in the bit stream file edit-compressed so that it may explain below.

[0060] When a user judges that video is received at Step 184 the already encoded

data is replaced by the data encoded newly. The following another section explains this step in more detail. If the data encoded newly replaces the data encoded before it will be judged at Step 192 whether a user wishes to edit further. In wishing to edit further more a flow returns to Step 174. If a user ends edit of video a flow will progress to Step 188 which ends an edited state.

[0061] When a user judges that video is not received at Step 184 a user ends edit operation at Step 186 so that edit change may not be recorded or a user reedits video. Video is reedited by the user who chooses a new parameter by edit. When a user wishes it is also possible to define the new video segment of an editing object.

[0062] If a user is satisfied with compression (it is one of the automatic-compression in automatic-compression or edit-compression) It will be in the state where preparation of the format of compressed video a compressed audio and each data file including all the other files to the last format was completed and will go into the format state 166. Information like the title memorized by the final formatted data file may be included in ancillary data. In the flow chart shown in drawing 6 with a block diagram. After the start of the format state in Step 200 if it determines that a user memorizes ancillary data the workstation 20 will read ancillary data in the digital memory storage 22 which possesses a floppy disk etc. in a floppy disk drive. And it is combined with the audio and video by which ancillary data was encoded at Step 204 and a final formatted file is made. The final digital bit stream created by a system can be built as a "program stream" defined as ISO/IEC 13818-1. Instead a final bit stream can be considered as the format of arbitrary kinds like the format of the kind of VBR (Variable Bit Rate) suitable for a direct satellite DSS format and use with an optical disc or others. Since the video and the audio which were encoded are separately memorizable the system can encode the same video and audio which were encoded to the separate last format. This is attained by the separate formatter using one formatter provided with the function which makes various formats. In Step 206 the formatted file is memorized by the disk 21.

[0063] The formatted file will be in a completion state at Step 208 written in the tape in the digital memory storage 22 of Step 210. It goes into exit status at Step 122 which performs various "housecleaning" functions to delete after that the temporary file which is not needed any longer. When an abnormal termination demand like [before a compression process is ended] occurs a cancel procedure is performed by workstation 30 which stops an audio tape player and a video tape player and carries out the cleanup also of the data file which may be destroyed. A process is ended at Step 216 from it.

IV. a system substitution measure -- although a desirable system component and its operation were explained above needless to say alternative hardware can substitute the hardware indicated by drawing 1 A and drawing 1 B. As for the video encoding device 50 when it can use it is desirable that it is MPEG-2 video encoder. However the ISO/IEC 1172-video standard encoder which is known also as MPEG-1 video encoder and which is marketed can also be used. An MPEG-1 buffer-management module needs to enable it to control the required

improvement to MPEG-1 encoder from external source like the workstation 10. Since ITU-R 601 (front CIR 601) resolution video is stored the input video support must increase. Such improvement can be performed by the person skilled in the art of digital video encoding of not conducting an excessive experiment either.

[0064] Although four audio tape players and four audio encoders are drawn on drawing 1 the other composition is easily employable as an audio system. For example one set of 8 mm digital audio player marketed can be used instead of two or more DAT players. When one set of a TASCAM digital audio player is used being needed from the workstation 40 is only one RS-422 control signal but a maximum of eight separate audio channels can be supported simultaneously. An audio encoder must be able to accept eight audio inputs simultaneously and it cannot be overemphasized that it must be improved in order that the audio interfacing unit 72 may also cope with the increase to eight from four of audio information streams.

[0065] As an audio and an alternative plan of simultaneous encoding of video an audio and video encoding are executed by separate time a location or its both and can be combined with a final format later. Use of a known time code like a SMPTE (United States movie television technology person association) time code format is needed for this. A time code is created by a video tape player a workstation or the separate stand-alone type time code preparing program.

[0066] Recording VTR 52 has an available video decoder and when connected to the workstation 20 it can be eliminated. In that case the video reconstructed will be created from the disk file in the disk 21 instead of being recorded from the video encoding device 50 in a compressed phase. Exclusion of recording VTR 52 will save the expense of a system greatly not only tape expense but in respect of device costs.

[0067] Commercial X-terminal option can be used for the capability to perform graphic operation in respect of graphic overlay of display video and it can support it on a console display. For example console DIPUREI 31 is replaced at X-terminal provided with the function which displays video in an overlay graphic window. A multimedia display terminal like HDSV View Station of Human Designed Systems of King of Prussia of Pennsylvania marketed It is connectable with X-terminal for a display and manual editing region definition operation. However the video of the quality of an occupation article of a video decoder must be expressed as a monitor for occupations like the monitor 61 illustrated by drawing 1 A so that the quality of the video signal with which human being's operator was reconstructed can be evaluated.

V. the number of audio encoders is reduced maintaining the optimal encoding efficiency -- as mentioned above drawing 1 A The block diagram of the encoding system provided with four sets of the multipass video encoding device 50 and the one pass audio encoders 70a 70b 70c and 70d is carried out. Playback VTR 51 supplies the video signal which should be encoded by the video encoding device 50 and digital audio tape (DAT) player 71a-71d supplies the audio signal which should be encoded by one pass audio encoder 70a-70d respectively.

[0068] In the example of this invention eight audio tracks in the last formatted audio

image data stream are encoded. Each audio track possesses one or more audio channels. Although the audio track can store an audio with other devices a certain audio track may be provided with the left channel and right channels for example. It explains to drawing 7. But the method of asking for the number of audio encoders in order to execute efficient audio image encoding and the method of executing this audio image encoding are explained to drawing 7. In drawing 7 several P of the path needed for video encoding at Step 240 is calculated after a start. In order to encode video to an MPEG format two or more paths are usually needed for encoding of video. In the case of this invention a desirable number of passes is 2 but three or more paths are also possible. And several T of the audio track encoded at Step 242 is calculated. As for the number of audio tracks eight is desirable although the number of audio tracks which is different in this invention is also possible. Next in Step 244 number AE of one pass audio encoders needed in order to execute most efficient audio image encoding is calculated. The number of required audio encoders is equal to the number of audio tracks by which division is done by the number of the paths needed for video encoding and which should be encoded. Eight audio tracks by which division is done with two paths of an encoder in an example are equal to four required audio encoders. When the number with which the audio encoder was calculated is not an integer it is necessary to revalue the number with which the audio encoder was calculated so that it may become an integer. For example when only seven audio tracks are encoded by the two pass video encoder it is set to $7/2=3.5$ and 3.5 is revalued by the integer 4 largest next. Between the encoding processes performed by this system although one video path uses only the inner 3 set of an audio encoder the other video encoding path uses four sets of all audio encoders. It should be cautious of it not being necessary to calculate the number of the one pass audio encoders needed whenever Step 240-244 needs to be performed only at the time of the preliminary design of a system and audio image work is encoded. A computer / machine mounting is carried out by computer like the workstation 30 or Steps 240, 242 and 244 are judged by the designer of a user or an audio image encoding system. [0069] After the number of audio encoders is called for and an audio image encoding system is set up an audio image encoding process can be started. A counter is set to 1 in Step 246. a counter -- each path of a video encode system -- counting -- business -- **** -- a sake -- it is variable. And although Step 248 is not all the audio tracks it performs the 1st video encoding path executing one pass audio encoding by some. For example when eight audio tracks and one two pass video encoder exist four audio encoders can be used during the 1st path. And a counter of 1 increases at Step 250. The following video encoding path is performed in Step 252 using the same audio encoder as having been used at Step 248 and executing one pass audio encoding of the audio track which is not encoded. For example four audio encoders used at Step 249 can encode the audio tracks 5-8 between the 2nd video encoding paths. And at Step 254 it is judged whether a counter is equal to P and an indispensable video encoding number of passes. In an example a process will be ended since there are only two video encoding paths.

When not ending a flow returns in order to execute the loop which comprises Steps 250, 252 and 254.

[0070] The other example but needless to say not only using the two pass video encoder for audio image works with eight audio tracks is also possible. For example a three path video encoder can be used by six audio tracks. In this situation the most efficient number of audio encoders is 3 ($6/2=3$). The most efficient number of audio encoders will be set to 3 when a two pass video encoder is used and a total of six audio tracks are encoded. As an alternative plan a three path video encoder can be used by a system with five audio tracks to encode. Only one of audio encoders is used in this alternative system while both audio encoders are used and it is generated by both of other two video encoding paths two one pass audio encoders in the state where any one of video encoding paths is operating are required. If the instruction written in this specification is taken into consideration combination other than this of a video encoder and the number of audio tracks is also possible so that it may become clear in the person skilled in the art of a Prior art.

[0071] With the important whole surface of this invention in the point of being as efficient as possible while audio encoding is always executed while being generated by the path of video encoding and audio encoding is performing I hear that the path of video encoding is always performed and it is. In this agreement between the one passes of video encode operation even if there are few audio encoders some will always be used. In order to be as efficient as possible when not doing division of the number of audio tracks uniformly by the number of audio encoders () [get blocked and] If division of the number of the audio tracks encoded is done with a video number of passes when remainder will come out there is an audio encoder which had stopped the number equal to the number which deducted remainder from the video path in the whole encoding process. For example in the case of a three path video encoder and four audio tracks two audio encoders are needed and 1 comes out not much (remainder is 1 for example at $4/3=1$). Therefore in order are the most efficient and three path video encoding is completed and to complete audio encoding simultaneously with **it is only 2 times that an audio encoder is stopped (for example when 1 is deducted not much from three video paths it is equal to 2). It can generate with the same path twice which an encoder stops (there are two audio encoders which are different on the same video encoding path) It means that there are two video encoding paths where each of the two video encoding paths with said other example utilizes all the audio encoders or which have one pause audio encoder. In these two paths the same or another audio encoder may be stopped.

VI. The log file for memorizing the information on the encoded video And the encoding system with which the log interpreter utility indication of [for accessing a log file and the encoded video] was carried out At the beginning video is encoded to compressed digital FOMATTO like MPEG-2 MPEG-1 or the other digital video encoding format. the [concerning / this encoding / a system component] -- the [about II paragraph and a general-purpose system action] -- it explains to an III

paragraph. Therefore the encoded video data is memorized for example within the digital format on a hard disk.

[0072] The statistics and the other information on video are created between digital video encoding processes and a log file memorizes. The person skilled in the art of a Prior art well versed in MPEG video encoding has full knowledge of the method of creating the statistics which describe video the kinds of information these statistics and information. In the conventional encoding system although this kind of information and data are created they are usually discarded after video encoding. However the sum total bit which in the case of this invention was used in order to code a picture This created information on the mean square error which encodes a picture the normal child-sized scale per picture etc. is dumped by the log file from system RAM after an encoding process. The example of the log file used by this invention is explained to drawing 8 A-10B. Especially these log files are effective in a VBR system and are created between the initial paths (there may be plurality) of a video encoder. And a log file is used between the edit between the last encoding paths and after video and re-encoding.

[0073] There are the log file of a sequence layer the picture layer log file and three kinds of different log files called the log file of a macro block layer which may be used by this invention. Each of these log files describes the statistics and the information on each of that layer. There is a log file format of two kinds of different macro block layers called the full format shown in drawing 10 (A) with a block diagram and the shortening format shown in drawing 10 (B) with a block diagram. When the detailed information on macro block layer ** is needed a full format is used and a shortening format is used when a follow-up survey of detailed all of macro blocks does not need to be conducted. Since the number of the macro blocks in the movie of 2 hours is a large number a macro block log file consumes a lot of memory spaces. Therefore unless a big memory space can be used the macro block file of completeness or a shortening format may not be memorizable. However when macro block information needs to be obtained encoding video can be decrypted or this information can be reconstructed by the other means.

[0074] Many of detailed log file formats shown in drawing 8 A-10B with a block diagram are not important and it should comment on these file formats to having been provided only as an example figure of the data which can be created between encoding processes. However since some of information in a log file is used between the processes of changing the quality of video they are important. In order to determine the edit point of the encoded video so that it may explain below it is necessary to sum up the number of bits consumed by the early picture. The information totaled is drawn on drawing 9 B and drawing 9 C and The number of bits for every picture (generated_bit_picture) A sequence header (bits_sequence_header) It is related with the number of bits (stuffing_size) used in order to fill the number of bits (bits_GOP_header) of the group (GOP) header of a picture and the number of bits used as stuffing or a space.

[0075] Drawing 11 explains the flow chart of various functions performed using a

log file. At Step 270 the log file of the picture created from automatic execution of the video encoder shown in drawing 8 A-9C with a block diagram and a sequence layer is read after initial encoding of a video segment. And the index table showing the byte offset of the log file record for every picture recorded on the video encoded at Step 272 is created. Two tables are created. As for another side an index is attached to one side with a time code by the frame number. The byte offset in the log file for every picture is remembered by the time code index table and it is rearranged into it in ascending order of a time code. The byte offset into the log file which similarly is rearranged into a picture frame index table in ascending order of a picture number inputted (it is checked by an encoder like) is memorized.

[0076] The created index table is effective in selecting quickly the log information of the arbitrary pictures of which it is expected or a macro block. In order to memorize the data of the sequence in a log file a picture or a macro block the fixed size record is used but an encoder rearranges a conversion picture before coding them. In addition when dealing with the movie (per second 24 frames) sauce pulled down by per second 30 frames at a rate of 3:2 a certain kind of time code is omitted. An index table serves as a means for tracing the position of suitable data quickly in spite of rearrangement of a picture and the abbreviation of a time code.

[0077] Since the encoded video changes frequently as the various frames and periods of video are manually changed in the quality The encoding system by this invention should comment on not maintaining the directory or index of the encoding number of bits used when coding each picture. However since the log file which describes encoding information is a fixed size record even if the quality of a picture or a period is changed it is easy to maintain the directory or index of information in a log file.

[0078] The user can obtain various information quickly by using Step 274 after creation of an index table from the log file which uses various log file utilities drawn on the blocks 276, 278 and 280. When the record in a log file is needed the created index is used in order to trace the record of the hope in a log file and a perfect record is pulled out from a log file. And this record is analyzed in order to search each parameter of which it is expected. The process of analyzing each record using an index table serves as all the foundation of other log interpretive program utilities. An analysis process is known and the person skilled in the art of a Prior art understands the method of retrieving the information on hope from the composed file.

[0079] Step 276 is a cut-and-paste utility of a log file. It is in a frame or after changing quantization (and image quality) manually in a period the original set of a log file must be updated so that it may correspond to change generated during edit. The cut-and-paste utility of a log file traces the position of offset within the log file corresponding to an edited segment time code and replaces the information in a log file for the new information corresponding to the re-encoded video. The process of re-encoding is explained to the following VIII paragraphs.

[0080] In Step 278 in order to enable edit of the bit stream encoded in digital

on the utility for a parameter to come to hand is shown. When video is encoded first and a user wishes the quality of the frame of video or change of a period it is deleted and it is necessary to trace the suitable portion in the encoded video which must be replaced by the video provided with the quality of hope encoded newly.

[0081] The byte offset of the edit point in the bit stream of the encoded video is called for by calculating the accumulation item of the bit used for encoding each picture to the picture corresponding to an edit point. In order to ask for the total of the byte from the beginning of a file or a bit in which the appointed picture is located the number of bits used for encoding each of a former picture is totaled. With the information which comes to hand from the totaled picture log file. They are the number of the bits B_i used for encoding the i -th picture the number of bit S_i which constitutes a sequence header the number of the bits G_i used for encoding the group (GOP) header of a picture and the number of bit T_i used as stuffing. Bit offset of the N th picture is calculated as shown below.

[0082]

Byte offset is calculated only by doing division of the bit offset by 8. In order to search for the address offset in the encoded video bit stream as an alternative plan which uses a log file a bit stream decoder can be used and the directory of picture offset can be created. However this approach is troublesome in order to have to perform the whole bit stream whenever edit finishes in order that an encoder may update a directory.

[0083] Finally Step 280 shows the utility for rounding off a time code to the beginning of the group (GOP) of the nearest picture or the last. All the cutoff or edits of the encoded bit stream are performed on the boundary of the group (GOP) of a picture. However the operator which performs edit in video does not understand a GOP boundary. Once an operator specifies the start points and the end point of an edit segment a system will calculate a GOP boundary. A GOP boundary is set up by examining the picture layer parameter called the GOP number memorized in each picture record. If the GOP number (GOP_number) in drawing 9 B changes by tracing through the continuous record of a picture layer log file it will be judged that there is a GOP boundary.

[0084] Since the log file and the log file utility are closely related to a video encoder a log file is memorized in the hard disk of the workstation 10 illustrated by drawing 1 A and a log file utility is performed by workstation 10. Instead a log file may be stored in the workstation 30 or the other arbitrary processing systems and a log file utility may be performed.

VII. Change A. of the quality of the encoded video The encoded video which is created by the change encoding system of the quality in a period aims at being written in a final digital storage medium like an optical disc. The encoded video on a disk is encoded with the Variable Bit Rate (VBR) which provides the possibility of the major change of the bit rate between the sections where videos differed as

compared with fixed bit rate (CBR) operation. Since the remarkably different number of bits can be assigned to a different picture now by VBR coding image quality changes temporally. Furthermore VBR also provides a possibility of reassigning a data rate to the section of the video defined by the user within an impossible scale in CBR operation. Probably a very big buffer is needed by a decoder when CBR attains the same thing.

[0085] Although it becomes possible to be called manual edit by an encoding system manual edit is not replacing a scene which shortened or extended the scene in the movie or is different but is changing image quality. Image quality can be changed temporally and it is called time edit or time manual edit so that it may explain to this paragraph. In another paragraph by shifting a bit from a section with a frame to another section of a frame explains being called the space edit or space manual edit which changes the image quality within the frame of data.

[0086] It is necessary to fill time edit, i.e. three main restrictions which show below the bit between the sections where edit segments are various by rediscussing reliance.

[0087] 1. Don't exceed the capacity to which the total capacity of the movie/video edited was assigned.

[0088] 2. There needs to be no buffer underflow into an edit segment. And there needs to be no buffer underflow in the outside of the edit segment produced by 3. time edit.

[0089] The restrictions 1 are processed by reencoding an edit segment so that the number of bits as the segment encoded first with same edited segment produced as a result may be consumed. It should comment on it being the whole period when a variety segment is examined. Quality is specified by the user and these periods of an edit segment are only called a section or a period to a different period in an edit segment.

[0090] In order to trace whether an underflow state exists it is required to check a target bit using the model of a decoder. In VBR operation of an encoder the version by which the video buffer verifier (VBR) model was corrected is used. This model is called the simple correction model of VBR and can be used for a digital memory storage medium (DSM). VBV for VBR modes is known and is defined by MPEG-2 standard and is indicated especially to the Appendix C of the C.3.2nd paragraphs of ISO/IEC 13818-2.

[0091] Although the buffer of a decoder is still full of the VBV model for VBR it is restricted so that an underflow may not be carried out. It is filled up with a buffer until a buffer fills with the fixed rate R_{max} and a bit is not transmitted to a buffer until a buffer begins to become empty. The bit needed in order to decrypt a certain frame is deleted from a buffer for every frame interval. When there are too many bits which will be assigned to the sequence of a picture in a VBR mode if it puts in bearing this in mind it turns out that a buffer underflow occurs. A buffer becomes empty and when deleted at a rate earlier than the rate to which the large majority of the bit of a high-definition picture can charge a buffer a buffer carries out the underflow of it as the bit of a picture is deleted from a buffer. In order to

prevent an underflow the number of bits used in order to encode a picture is reduced. If the number of bits of a picture is reduced a restoration rate will be kept the same but the rate which deletes a bit from a buffer is reduced. Since bit stream transmission of VBV buffer HE stops immediately when VBV occupancy is full buffer overflow cannot be generated about VBR. That is not buffer overflow but a situation only full of a buffer exists. It is like and as [being explained to ISO/IEC 13818-2 Appendix C and the C.3.2nd] follows [model / of VBR / simple correction VBV].

[0092] 1. bit_rate_field in sequence_header expresses max_bit_rate.

[0093] 2. VBV is empty at the beginning. After filling a VBV buffer (that is VBV buffer file) to the limit [the rate of max_bit_rate] a decryption process begins. Picture data is immediately deleted and encoded like the VBV model of CBR defined as MPEG-2 international standard.

[0094] 3. A bit stream is stored in a VBV buffer at the rate of max_bit_rate until a VBV buffer fills. When VBV buffer occupancy is full the bit stream transmission to VBV stops immediately. After the data of the following picture is deleted bit stream transmission is started immediately.

[0095] 4. The underflow of the VBV buffer shall not be carried out.

[0096] 5. Each vbv_delay field is set as FFFF.

[0097] In order to change the quality of video in a period it is necessary to delete a bit for a bit from the frame of an addition or video on the frame of video.

Here explanation about the number of the bits deleted from the section of an addition or video into the section of video is given. It depends for the number of the bits which can be deleted from the section of an addition or video into the section of video on the available bit rate of the section of the video currently examined. Available bit rate R is calculated as follows.

$R = (\text{the total quota bit} / \text{frame number}) * \text{Frame_Rate}$ Since the (2) total quota bit is in direct proportion to the bit rate referring to an assigned bit number tells the same meaning as the bit rate except for the case where different units are used. When the guaranteed minimum rate for example R_{\min} assumes that it is applied to the arbitrary sections of video the maximum number of bits which can be deleted from a section "i" is as follows.

$T_i = (R_i - R_{\min}) * N_i / \text{Frame_Rate}$ (3) In this case N_i is a frame number of a section "i" and R_i is an effective rate of the same section. The guaranteed minimum is needed in order to guarantee the minimum quality level to a section. The maximum number of bits which similarly can be added to the specified section "j" is calculated as follows.

$T_j = (R_{\max} - R_j) * N_j / \text{Frame_Rate}$ (4) R_{\max} is a rate with which it is filled up when a maximum rate i.e. a buffer is not full in this case and N_j is a frame number of the section j.

[0098] Explanation about changing image quality in a period (time edit) is shortly given about the figure referred to drawing 12 and there. after a start and Step 281 -- the -- the [II paragraph and] -- as mentioned above by the III paragraph automatic video encoding is executed without a user's intervention. In

automatic encoding there are the maximum rate and the minimum rate of a data stream which are created. A maximum bit rate is determined by the function of a video decoder for example R_{max} is Mbits per second. According to the minimum image quality of which it is expected a user can specify the minimum bit rate. The typical minimum bit rate R_{min} is per second 3 Mbits. It is desirable to set up R_{min} which is a little less than an average encoding rate unlike an average encoding rate so that it becomes the bit rate with a low portion of video and can become a video rate with a high portion of video. If R_{min} is slightly less than the average encoding rate of video and is set up the encoding video produced as a result will serve as considerably fixed image quality. Usually when R_{min} is abbreviation 3 Mbits per second in R_{max} = per second 8 Mbits the average value of video is per second about 3. It should be set to 5 Mbits from 1/2.

[0099] After move video encoding is executed at Step 281 it is necessary to search for a rate quantization function. Although the number of bits required to acquire the quality of hope is called for in the process of this invention since the bit rate of hope is produced as a result a quantized value must be set up at the time of re-encoding. The report with which a rate quantization function is taken in by this specification by reference respectively The IEEE report about image processing the 3rd volume No. 5 September 1994 533 – 545 – page Lamb Chan Derain (Ramchandran) -- "bit assignment for the dependence quantization applied to multi-resolution and MPEG video disk Rozsa" depended in addition to this. IEEE ICASSP1993 and V381-V-384 bit assignment for the dependence quantization applied to an MPEG video encoder by Lamb Chan Derain and others And it is calculable by the conventional method which is explained to "multiplexed common multi-channel bit rate control for MPEG videos which uses a rate distortion curve" by Leininger (Reininger) shown by the Princeton video conferencing in 1994.

[0100] Next in Step 284 the user quality preset value of the selected period is inputted. A user inputs ultimately the priority of a different section used for controlling quantization of a picture. Quantization controls the image quality called the resolution of MPEG encoding with a loss. In order to change the quality of video in a period the user INTA phase used in order to collect information is explained to drawing 13. At the beginning a user inputs the period of the whole edit segment which will change quality. Although not illustrated by the user interface of drawing 13 the input of this retribution is inputted before the user interface of drawing 13 is displayed. In the case of the example shown in drawing 13 the period of the edit segment of a change target may be from time 0 minute for example till time 5 minutes. After a user describes the total period of an edit segment various priorities of the ranges from -5 to +5 including the priority P are specified as the still shorter period in an edit segment the quality of this period is protected and it is shown that it must not change. The priority "nothing" is the same meaning as the priority P. In order that an encoding system may not be limited to said input scheme but may input the priority of a user's hope it cannot be overemphasized that the other range or sign can be used.

[0101] In addition to specifying a priority a user also inputs the minimum bit rate.

This bit rate may differ from the minimum bit rate used in order to encode video between automatic video encodings of Step 281. Another feature of a user interface is the point that a user can set up a "background priority" when it is used. This means that this background priority is set as the period within the chosen period when a user did not specify a priority. For example in drawing 13 supposing "1" is specified as a background priority even if a user will not specify a priority as this specific section manually from the unspecified period 00:03:20:01 00:03:59:99 the priority "1" is specified automatically.

[0102] In the pars basilaris ossis occipitalis of the interface drawn on drawing 13. In order to create a new period within an edit segment insertion block is used in order to change the period created before a "recall" block is used since change added to a period is saved updating is used and "deletion" is used in order to delete a period. An encoded ultimate digital memory storage medium for videos like an optical disc is provided with the restricted storage area capacity which is ideally filled to capacity when the data encoded automatically is written in a disk. Therefore image quality is raised and when pulling up the bit rate of a certain specific period in order to supply an excessive storage capacity always required for the period whose quality improved the section of video to which quality fell must exist. Therefore it is impossible to assign high image quality only in one period without assigning the image quality which is different in another period in order to have to take a bit required in the example of this invention to improve quality from another section of video. Therefore unlike a background section in order that the time manual edit may operate appropriately a section "with protection" and at least one section which is not marked must be in the segment of the edited video. As a result of a user's setting up a priority the same image quality and bit rate So that it may become clear if you understand the process for not producing whenever they are chosen but asking for the number with which the bit of the picture in an edit segment was corrected It depends for the quality of a certain specific section also on the length of the priority and other sections where other sections were assigned.

[0103] the feature utilized by the encoding system by this invention -- a televiewer -- the inside of an automatic encoding process -- or it is the point that the encoded video can be reviewed later. As the encoded video is displayed on a televiewer a televiewer Even if video is not suspended or it does not forget to write an interested time code in order to memorize interested point and field of video functions such as pressing the return key or "clicking" etc. a graphic user interface on the "button" displayed on a screen can be performed. A user calls the point with such concern and even if it does not review the whole movie it can be considered minutely later. This serves as the convenient and efficient way a user determines the point in the encoding video which wishes to change. This function is attained by memorizing a time code corresponding to the position in the time of interested point and field occurring.

[0104] After a user inputs the quality preset value of hope at Step 284 the bit corrected for every frame based on a user's input at Step 286 is calculated. This

process is described in detail about drawing 14–16B. The general procedure for changing the bit rate of time within the edit segment of video is explained to the flow chart of drawing 14. In order to calculate the bit rate based on a user's input in Step 302 the input of the user who suited within the limits of between 10 and 0 from within the limits between –5 and +5 is mapped. –Correspond to the mapped preset value 10 and it is the minimum quality an input called +5 corresponds to the mapped preset value 0 and the initial input 5 is the maximum image quality. Similarly the inputted quality selection which is called 0 is corresponding to the preset value mapped [which it is called 5] etc. multiplication is carried out to the purpose of mapping performed at Step 302 by the preset value by which the assigned number of bits was mapped -- it states quantity –1 and it comes and is proportional to e (2.71828 ...) which can be pulled up until. In this case as the number of bits when the preset value of a priority is still smaller is less from the case where a priority is high further e^{-10} are less than e^0 . The process of asking for a target bit corresponding to priority –4 to +4 is explained to drawing 16 A and drawing 16 B in detail.

[0105] It is judged in Step 304 whether there is any section "with protection." The section with protection is a section which must not change image quality from former automatic encoding. When there is a section with protection a flow progresses to Step 306 and is set up become equal to the number of bits before the number of bits assigned to re-encoding of a section with protection there arises as a result of the first automatic encoding. Thus the picture quality of a section with protection does not change.

[0106] In Step 308 it is judged whether there is any section set as the maximum quality or the minimum quality. With the maximum quality it corresponds to the user input +5 and corresponds to user input –5 with the minimum quality. In the case of the maximum image quality and the minimum image quality it is desirable for the encoded video by which the video of the top quality and the video of the minimum quality were set up respectively to come to hand without exceeding the storage capacity of the medium which still fills restrictions of a video buffer and memorizes the encoded video. When there is a section set as the maximum quality or the minimum quality a flow progresses to Step 310 and the correction number of bits of the maximum case or the minimum case is calculated there. This process is shown in drawing 15 in detail.

[0107] At drawing 15 the process of asking for the number of bits assigned to the maximum case and the minimum case begins at Step 320 by assigning the minimum rate and a maximum rate to the section of the video assigned to 10 and 0 of the mapped priority respectively. In Step 302 of drawing 14 as long as the remaining portion of description of this mapping and time edit is performed and there are no other clear directions the mapped priority is referred to. The maximum data rate receives restriction with the capability of a decoder serves as per second 1 Mbits and so that it may be described below the minimum data rate A block diagram is carried out by the user interface by drawing 13 as shown and if it may be less than the minimum bit rate used between the automatic encoding modes

which usually serve as per second 3 Mbits(es)it may exceed.

[0108]And in Step 322it is judged whether there is any section where the mapped priorities 1–9 were set up. That isin addition to the maximum quality or the minimum quality being assignedthe other quality of an intermediate item which a user assigns exists. When there is the other mapped prioritya flow progresses to Step 324 and the average bit rate of all the sections where the mapped priority between 1 and 9 is set up is calculated. When this average is not in tolerance level (between a maximum bit rate and the minimum bit rates)a flow progresses to Step 328. In Step 328a maximum ratei.e.the rate which realizes the highest image qualityis reducedwhen an average is less than the minimum rate. By reducing a maximum ratethe number of bits which can be used for the non–maximum situation and the non–minimum situation (mapping priority 1–9) increasesandtherebyan average becomes within the limits of hopeor near the range of hope. Insteadin Step 328when an average exceeds a maximum ratethe minimum data rate can be pulled up andtherebythe number of bits of the non–maximum situation and the non–minimum situation (for examplepriority 1–9) is reduced. After Step 328a flow returns to 324 by which Steps 324326and 329 are repeated until an average becomes within the limits. A maximum bit rate and the minimum bit rate are adjusted in order that a section “with protection” may ensure receiving the same bit as adjustment before of quality.

[0109]If it judges that within the limits once has an average at Step 326a flow will progress to Step 330 and the minimum rate and maximum rate which were calculated by the section into which the priorities 10 and 0 mappedrespectively are set up will be assigned. And the remaining bits are distributed to the section which is not in protectionand the maximum situation or the minimum situationeither at Step 332. The section of video and the changing method of a data rate with which the mapped range 1–9 is set up are explained about drawing 16 A and drawing 16 B later.

[0110]It is Step 340 when the section where the priority 1–9 mapped at Step 322 is set up did not existi.e.it is judged that only the maximum sectionthe minimum sectionor a section with protection existsThe minimum rate and a maximum rate are adjusted if needed until the last assignment of a bit becomes equal to initial assignment of the bit by automatic encoding. There are two kinds of alternate methods which perform Step 340. Maximum bit rate R_{max} is being fixed to the 1stfor example by per second 8 Mbits. R_{min} which is the minimum data rate since the sum total bit of the synthetic edit segment is understood is only adjusted in order to attain the same bit totalif gained between automatic encodings. When R_{min} cannot be adjusted in order to attain the number before the total bitin order to attain the bit total gained between automatic encodings,the maximum data rate R_{max} is adjusted.

[0111]In order to specify the priority 9 as the priority 1 and the minimum quality section into the maximum quality section at Step 340 and to ask for the new bit rate of the period of an edit segment as an alternative planthe bit quota algorithm shown in drawing 16 A and drawing 16 B with a block diagram is performed. And it

is assigned to the section into which the minimum specification quality and the highest specification quality are set up and the process of drawing 15 is completed so that the data rate calculated according to drawing 16 A and drawing 16 B at Step 342 may explain below.

[0112] When Step 308 and Step 310 are required after they are performed it is judged whether the priority 1–9 mapped by all the sections is set up at Step 312. If this is applied a flow will progress to Step 314 and the correction number of bits of the mapped priority will be calculated so that it may explain in full detail to drawing 16 A and drawing 16 B.

[0113] The size of the common bit pool used is determined in the step of the beginning of drawing 16 A. At the beginning all the section **** to which the mapping priority 1–9 which it is not with protection is set and a fixed number of bits are deleted. For this reason the coefficient k_1 called redistribution depth is used for example is set as 0.440% of the bits assigned before (from jam automatic encoding execution) are not with protection either and it means being deleted from all the sections that are not even if marked by the maximum quality and the minimum quality.

[0114] Next in Step 362 it is judged how much user choice quality affects quality. The variable k_3 called a dynamic range coefficient is set up and it opts for quality change which will be produced from a priority. There may be [which imposes change of various quality like big improvement which is comparatively alike in small improvement or quality] the priority 2 relatively in quality corresponding to how k_3 is set up. Usually the dynamic range coefficient k_3 is set as 1.0.

[0115] In order to use an user choice priority by redistribution of a bit the exponential relation of the form $\exp(-\text{priority} / k_3)$ is used. The dynamic range coefficient k_3 in an index determines the shape of an exponential curve. A curve serves as a steep slope so that the value of a dynamic range coefficient becomes large and the imbalance in the bit assigned and done again between the sections where the changing priority was set up on parenchyma becomes strong. The bit assigned to that section decreases that is the mapping priority 1 is the highest priority as the absolute value of the mapped priority increases this negative index (still lower priority section) It means that the priority which will be set up if a number is pulled up to 9 becomes low.

[0116] An user choice priority is normalized in Step 364. That is the priority which the user chose is used in order to determine the relative priority of each section. First E_1 is calculated.

In this case p_1 is the mapped user choice priority and i is a section period in an edit segment excluding the period when protection the maximum quality or the minimum quality is set up. Next the product opportunity-ized coefficient k_2 made from a priority is calculated as follows.

[0117]

In this case N is the number of sections. A priority normalization coefficient is needed in order that the sum total of all the bits assigned and done again according to the priority may ensure always becoming equal to the usable common pool of a bit.

[0118] Therefore the relative priority of each section is calculated as follows.

[0119]

Since R_i is a fraction of the relative priority of the i -th section the summation of all the R_i (s) becomes single.

[0120] The size of a section is normalized at Step 366. Since the sizes of the period when the assigned priority is set up differ and a bit must be distributed in proportion to the size of various sections this is required. The 1st is asked for fraction T_i (the non-maximum or the frame to which the non-minimum priority is set which it is not with protection) of the frame number within each period in an edited frame total according to the following.

[0121] $T_i = N_i / N$ (8) In this case N_i is a frame number within the i -th period and N is a frame total of an edit segment. That is they are all the frames as which it is not changed and protected but the priority which is not a priority which is not the maximum or the minimum either is specified. Therefore k_3 size regular family lines are calculated according to the following equations.

[0122]

In this case N is the number of sections. A flow progresses to Step 368 of drawing 16 B from Step 366 of drawing 16 A.

[0123] At drawing 16 B the common pool of a bit is called for at Step 368 according to the following equations.

[0124]

In this case B_i is a bit total adding all the bits distributed to the number of bits used between the automatic encoding procedures of the i -th section at Step 332 of drawing 15. Summation is on the " N " section.

[0125] And it asks for the number of targets of the bit assigned to each section at Step 370 according to the following equations.

The bit total F_i of each section is equal to the thing adding the number of original of the bit (B_i) which was given to the percentage of the common pool of a bit at the common pool of the bit and which was reduced by the bit. The percentage of the bit from the common pool of a bit the frame summed up by the fraction (T_i) of the frame total of each section the relative priority (R_i) of a section and each of a

section -- it is based on the size normalization coefficient (k_3) depending on the priority of the section by which multiplication is carried out in the corresponding fraction.

[0126] There may be no number of targets of the bit which was calculated at Step 370 and which was assigned for every section between the maximum data rate and the minimum data rate. If this is applied at Step 372 the redistribution depth k_1 is reduced 0.01 times and a flow will return in order to recalculate the common pool of the bit of Step 368 and to recalculate the number of targets of the bit of Step 370. It performs until the number of targets of the bit by which the loop of Steps 368, 370, 372 and 374 was calculated at Step 370 comes between the maximum data rate and the minimum data rate.

[0127] And in Step 376 bit scale factor S_i intra-frame one and non-intra-frame one need the number of bits which is different in order to acquire the quality of hope. First bit scale factor S_i is calculated according to the following equations.

Next the scale factor of an intra picture and a non-intra picture is calculated. At least intra is taken and as for the factor S_i equation ***** calculation of the following is done.

By big intra-frame one a buffer underflow may occur and the Intra factor I_f By restricting change by bit assignment of an intra picture to 20% at the time of $I_f = 0.8$ shows a safety ratio (it is got blocked and S_i is restricted to range $0.8 < S_i < 1.2$). It is cautious of it being $S_i' - S_i - 1.0$. The non-Intra bit scale factor S_{Ni} is calculated as follows.

In this case B_{Ii} and B_{Ni} are the total Intra bits and non-Intra bits which are totaled with the i -th section by a picture type.

[0128] Finally the number of bits used for every frame is called for using the number of bits of the suitable origin by which receives and multiplication is carried out by a factor (either the Intra scale factor or the non-Intra scale factor). And the process of drawing 16 B (and drawing 14) is completed.

[0129] At this time the number of bits with which each frame was corrected is calculated and Step 286 of drawing 12 is completed. However a buffer underflow may arise with the changed number of bits of a frame or occupancy of the number of bits in the falling end of an edit segment may change so that the frame after it besides an edit segment may carry out an underflow. If these situations are necessary they will be investigated and corrected at Step 288 of drawing 12 explained by drawing 17 in full detail.

[0130] Drawing 17 explains the process of checking [whether there is any buffer underflow and] correcting and checking and correcting the problem for the bit rediscout reliance in the falling edge of an edit segment. At drawing 17 each bit

assignment of the frame of an edit segment is scanned at Step 390. And it is checked at Step 392 whether there is any underflow of distribution of the new bit which uses the above-mentioned VBV model. An underflow problem exists (u_flow list) the number of bits required to eliminate each underflow is called for and the list of points memorized by u_val is created. It has on its mind that an underflow occurs for the picture expressed by the bit which are and Lycium chinense is important. When the quality picture expressed with many bits is read from a buffer by the decoder when a picture is decrypted many bits are deleted from a buffer and these bits cannot be filled again quickly enough. Therefore in order to eliminate an underflow picture sky deletion of the bit is carried out. In Step 392 the list of points full of a buffer is created and these points are memorized by o_flow. Since data will no longer be transmitted to a buffer if a buffer fills as mentioned above even if a buffer is full it is cautious of it being satisfactory.

[0131] In Step 394 buffer occupancy in the falling edge of an edit segment is investigated. In a VBR mode of operation buffer occupancy in the falling edge of an edited video segment may differ from buffer occupancy of the corrected execution even if the edited segment is replaced with the completely same number of bits. This is produced from the singularity of the VBV model for VBR. Therefore it is important to investigate restrictions of buffer occupancy. Buffer occupancy in the falling edge of an edit segment is completely the same as redistribution before which is a bit or it is quite easy to show a certain thing [that are and a problem does not exist when still higher]. When buffer occupancy is completely the same as before occupancy of the frame after it also becomes completely the same as as with the execution before [which does not have violation of VBV / confirmed for coming out] lending. When occupancy is higher than before a buffer may come between the frames after it and when it is full of a buffer data is no longer sent to a buffer and since it does not generate an underflow is in a thoroughly permissible state. However as compared with former execution buffer occupancy falls and when still lower on the edge the 3rd case arises. With the frame after it this may create a buffer underflow and needs to investigate it closely. There are two kinds of methods of processing the frame of the last of an edit segment in order to prevent the buffer underflow of the frame in the outside of a buffer segment. The 1st approach is the easier one among two. With the number of bits of the frame of the last of an edit segment when larger or buffer occupancy equal to the original encoding than the original (a log file memorizes) encoding arises an underflow problem does not exist therefore does not need to devise treatment. When the number of bits in the buffer of the frame of the last of an edit segment is less than the number of bits in the buffer of the segment encoded first with this reduced number of bits. It may not be generated if the underflow of the frame of the outside of an edit segment may arise according to the contents of the frame after an edit segment. Instead of asking for the number of bits which is needed for protecting so that an underflow may not occur occupancy of the buffer of the segment of the last of an edit segment is only adjusted so that it may become the same as the case of the original encoding.

Buffer occupancy or sufficiency of the last of an edit segment can be pulled up from the purpose of safety after 10% of occupancy of the original encoding in order to guarantee that there is no possibility of an underflow. When it is necessary to pull up buffer occupancy of the frame of the last of an edit segment so that it may be set to the level of the original encoding as the number of bits of the end frame was performed at Step 398 it is necessary to decrease. The last frame is processed as an underflow point and it is stored in a u_flow list in order that even the level of the original encoding may pull up a buffer level the number of bits which needs to be deleted from the last frame is added to u_val .

[0132] The 2nd approach for processing the underflow produced by re-encoding of an edit segment is an iterative process which estimates the minimum possible level of the buffer of the frame of the last of the edit segment which does not cause a buffer underflow situation on the outside of an editing frame. This is performed by the process of estimating the maximum difference between buffer occupancy of the original encoding and buffer occupancy of re-encoding of the frame of the last of an edit segment.

[0133] When the flow chart shown in drawing 18 with a block diagram was seen B_0 was assigned to the original encoding buffer occupancy and the log file of the original encoding memorized. $i = 1 \dots$ Buffer occupancy with the frame " i " which exceeded the boundary of the edit segment in n is B_i . In this case i is equivalent to the frame of the last in the movie encoded. Whether it can become how much low from the original encoding without buffer occupancy of re-encoding causing an underflow situation by one of the continuous frames. It must be determined in contrast with the 1st approach that sets the buffer level of re-encoding as the point which was because it is the original encoding at least. New and also low buffer occupancy on the boundary which fulfills this condition is made into B_0' . This is the minimum permission buffer occupancy in an edit segment boundary. Since a difference is absorbed with the later frame interval after an edit segment a possibility of setting up buffer occupancy lower than buffer occupancy of former execution is produced from the fact that an underflow does not occur.

[0134] It starts at Step 420 of drawing 18 and buffer occupancy B_0' of the video by which the frame of the last of an edit segment was re-encoded assumes that it is zero without causing an underflow in the future. This must have higher buffer occupancy in order to prevent an underflow but since the buffer occupancy is not less than zero it is the free estimate about B_0' . However until value B_0' which does not cause an underflow is found in the process of drawing 18 whether an underflow arises by occupancy of the zero of the frame of the last of a re-encoding edit segment and when producing and it checks in order to check whether buffer occupancy is made to increase at the last of an edit segment.

[0135] The process of drawing 18 estimates the rough buffer occupancy difference " X " between the original encoding and re-encoding.

[0136]

"X" is checked and since the difference "X" cannot be determined eventually it is called estimate until being adjusted until all the possibilities of an underflow are eliminated if needed is guaranteed. It starts with the frame just behind a boundary frame i.e. $i = 1$ and the procedure of drawing 18 progresses one frame at a time at once. In the case of the 1st frame equal to zero B_0' is a value of the 1st-frame buffer occupancy of the original encoding as a rough difference is calculated at Step 422.

[0137]

Next it asks for the number of bits absorbed by the frame in the right-hand side of a boundary frame in Step 424. With the number of bits absorbed although it may be stored in a buffer by the frame after the edit segment of the original encoding it is an expression used when pointing out the bit which was not stored in the buffer between the original encoding since the buffer was full and a bit was not able to be accepted further actually. Although quantity $**ti$ was full of the buffer to the meantime since the value of B_0' was reduced it expresses now the period of the i -th frame of encoding of the origin of an after [the edit segment which can accept an additional bit]. It cannot become full [a buffer] of the whole frame but it is larger than zero or since it is equal to zero period $**ti$ must be less than the period of a frame. Between the original encodings when once is not full of a buffer either period $**ti$ is zero. It follows

In this case 'P' is measured per frame per second and is usually a picture rate which is per second 30 frames. As for a result full of a buffer the bit total absorbed on the right-hand side of an edit segment is calculated as follows.

[0138]

In this case R_{max} is a restoration rate of a buffer.

[0139] In Step 426 it is judged whether the procedure for asking for the buffer occupied level of the frame of the last of an edit segment can be ended. The sum total of the bit absorbed by the i -th frame exceeds X' that is since a buffer can be filled with a later frame so that Step 428 may be performed without a buffer underflow's occurring when it is $a_i > X'$ it is stopped by the procedure. Instead if the last frame n is reached ($I = n$) an underflow problem will not arise with the rough value of X' to the last of video.

[0140] When the reply to Step 426 is "no" the number of bits b_i which remains without being absorbed at Step 430 is calculated.

[0141]

And in Step 432 it is judged whether it is $b_i > B_i$. In this case it is buffer occupancy of the i -th frame after an edit segment in B_i . Since it remains without in $b_i > B_i$

absorbing many bits rather than the present buffer occupancy " B_i " approve the rough value of a difference and X' are too large and produce an underflow. Therefore only the " B_i " bit can absorb between the bits which remain without being absorbed without causing an underflow. Next rough X' of a difference must be revised at Step 434. In Step 434 the estimate at which X' was revised is calculated as a thing adding the bit absorbed by the bit which remains without being absorbed. [0142]

At Step 432 since buffer occupancy with the i -th frame is larger than the bit which remains without being absorbed when judging that it is $B_i < B$ there is no problem of an underflow it is not necessary to change the estimate of X' and a flow progresses to Step 436. In Step 436 the following frame is investigated (the increment of the i is carried out by 1) and it is repeated until the reply with a loop of Steps 424 426 430 432 434 and 436 affirmative to judgment at Step 426 is obtained. [0143] Since rough X' will be accepted if an affirmative reply is obtained by Step 426 in Step 428 minimum permission buffer occupancy B_0' in the boundary of an edit segment is as follows.

When actual buffer occupancy of the frame of the last by which the edit segment was re-encoded is less than B_0' it generates for the rediscount reliance whose underflow situation is a bit. In this case the last frame is marked as an underflow point in a u_flow list. The number of bits which needs to be deleted from the frame of the last of the edit segment for preventing an underflow situation (the number of bits of changed video and difference between number-of-bits B_0' which produces the minimum permission buffer occupancy as a result) It is stored in u_val the same with Step 392 having been described.

[0144] In the case of the encoding system of this invention it is desirable to know buffer occupancy of all the points in video and it should annotate to this information being memorized by the log file. However in reducing buffer occupancy with the frame of the last of an edit segment. It will be necessary to devise the corrective action which buffer occupancy of the frame after an edit segment is changed and the buffer occupation information in a log file becomes inaccurate therefore recalculates buffer occupancy of the frame after an edit segment.

[0145] If it is returned and referred to in the process of drawing 17 in Step 396 it will be judged whether an entry is in a u_flow list. When there is nothing an underflow problem does not exist but the process of drawing 17 is ended. When a u_flow list has an entry a flow progresses to Step 398 the section from the left-hand side of each position in a u_flow list to the nearest o_flow point is marked with protection and a bit equal to the corresponding number of bits in u_val is deleted from these sections. An underflow problem will be solved if the number of bits in a frame is reduced since a buffer underflow is caused by the bit in one or more

frames which are. In order to keep the fall with a bit remarkable in image quality owing to deleted in order to prevent an underflow from arising a bit is not deleted only from an underflow point but a buffer returns to the point of the full last and it is uniformly deleted from all the frames. This serves as the method of gaining the video of the top quality though an underflow is prevented.

[0146] Then these bits deleted at Step 398 are uniformly distributed to protection and the section which is not marked at Step 400 and the process of drawing 17 starts them once again. Since it does not become the guarantee of an underflow being no longer a problem even if it distributes a bit at Step 400 it is necessary to carry out repeated use of the whole process of drawing 17 and to check again whether there is any underflow situation.

[0147] After the process of drawing 17 is completed Step 288 of drawing 12 is completed and Step 290 of drawing 12 is performed. In Step 290 the normal child-ized value for every frame is calculated using the rate quantization function searched for at Step 282. Checked in order to check whether the number of the bits of which it is expected is before calculated at Step 286 and an underflow problem exists at Step 288 but. A quantized value must be calculated in order to re-encode data so that the number of bits of hope may be set as the encoded data. This searches for a rate quantization function by discovering a rate only calculates a quantized value and is determined.

[0148] When asking for a quantization level by a picture unit the rate quantization function (RQF) was searched for at Step 282. In order to estimate this function two data points are needed for every picture. The RQF is as follows.

[0149]

In this case a quantizing scale X_g and H of the number of bits and Q which are used in order that R may code one picture are the coding characteristics parameters of a picture unit. A quantizing scale expresses the independent bit so that H may become a bit a movement vector a DC coefficient etc. with which H is consumed by a header.

[0150] Execution of the beginning of an encoder can be used in order to create two data points for every coded picture in a video sequence. This can be performed with 1 set of mutual quantizing scales used for every macro block in a picture. These are processed in order to calculate a rate quantization model parameter " X " and " g ." These two parameters are memorizable in one and a file for every frame of the video coded with the value of H .

[0151] In order to ask for the quantizing scale which makes the target bit assigned to the coded specific frame with video said equation to which " R " expresses a target bit is used and the value of $(x_g H)$ is read in the memorized data file which may be the aforementioned log file. The value of the optimal " Q " for this equation is a quantizing scale of hope.

[0152] And the section where video was changed at Step 292 using the quantized value calculated at Step 290 is re-encoded. Unless a procedure special between

re-encodings is devised it should annotate to a problem arising by the re-encoding segment of the already encoded video by encoding. The details of the re-encoding process for avoiding encoding structure and decryption structure are explained to the following separate paragraphs. After video is re-encoded it is substituted for the bit stream of the video encoded before and a process is ended.

[0153] In order to judge whether video is permissible the user can display video after video is re-encoded so that drawing 5 may be described. When video cannot be permitted the user does not need to receive the video encoded newly and can continue use of the already encoded video. If a user understands that the video encoded newly is desirable the video already encoded by the video encoded newly will be replaced. And the user can edit the new segment of video spatially or in time so that it may explain to drawing 5.

[0154] Drawing 19 (A) and drawing 19 (B) show two different examples of the bit rate after the user from original video encoding inputs a manual quality priority. In drawing 19 (A) and drawing 19 (B) the bit rate (on 106 scales showing a megabit) is plotted to the group number of a picture. In the example shown with a block diagram the picture of 15 is stored in one GOP and each GOP is displayed every 0.5 second. A solid line shows the original encoding and a dashed line shows encoding after a user adjusts image quality manually.

[0155] Drawing 19 (A) explains encoding with five kinds to which it is set the priority 2-33 and 0 and with protection of different priority fields. As compared with the original encoding much more bits are used for the period when the priority 3 is set up after re-encoding. Since the priority is only merely only correlated mutually the bit rate of the section where the priority 2 is set up is not going up remarkably but the bit rate is falling with the section of -3 and 0. The bit rate after having existed before re-encoding and similar re-encoding must be set to a section with protection.

[0156] Drawing 19 (B) is the 2nd example with the field where the priorities 0204 and 0 which the user chooses respectively were set up. Here since the priority is correlated mutually it is a priority which does not mean that the bit rate has not changed the priority 0 but only faces the other section into which the priorities 2 and 4 are set up. In order to specify the high bit rate as the section ***** (ed) by the priority 4 these sections where the user choice priority 0 is set up are reduced and the almost same bit rate as encoding order is set to the section into which the priority 2 is set up.

[0157] Processing of time manual edit is generated by workstation 30 of drawing 1 A of the example of this invention. However since time manual edit is not generated between video encoding processes the workstation 10 can perform time manual edit calculation without reducing the rate of a video encoding process.

[0158] This invention should annotate to being described in detail as change of the image quality of various time segments within an edit segment. In order to permit change of the quality of the section of the picture which is not into the needless to say and same time segment it is clearly possible to extend this concept. For example in order to raise the quality for 5 minutes of the last of a movie at the

sacrifice of for 5 minutes of the beginning instruction of this invention is applicable. When being expected of changing the quality of the period on the outside of an independent edit segment so that it may explain below a user. An interested segment is displayed continuously and is checked a bit is reassigned and probably not only buffer occupancy in the falling edge of each section but the restrictions related to the sum total bit of a movie need to check that it is in a regular limit.

[0159] Time manual edit has been described as making the number of bits of encoding of the origin of an edit segment the same as the number of bits of re-encoding of an edit segment. However when the bit which a re-encoding edit segment consumes is further lessened when you wish or sufficient storage capacity exists the number of bits of a re-encoding edit segment may exceed the number of bits of the original encoding. The original encoding of video can also be executed so that no memory spaces on the last digital storage medium which can be used may be consumed by that cause. Therefore there are many bit totals after re-encoding between 0% containing the arbitrary quantity which exceeds or is less than origin for example few [that it is 1% as more as 5% or] bits and 20% or they can consume a small quantity.

[0160] Explanation of time manual edit has been given about changing video after video encoding. However the instruction written in this specification is applicable also to a system without an initial encoding process. Instruction of the time manual edit about use of a common bit pool is applicable to the concept of space manual edit of the video frame shown below. The form of operation in each frame of space manual edit is applicable to the sequence of a frame in order to attain high-quality time edit so that it may explain below.

B. Explain the general procedure for changing the quality of a field into the changed drawing 20 of the quality in the field of a frame within one or more frames. After a start at Step 450 automatic video encoding which encodes input video to a compressed digital format like an MPEG format or the other format is executed so that it may explain to the paragraph about a system component and a general system action. In Step 452 the video encoded by the user is displayed and the user can input the command which shows change of the quality of the field in one or more frames at Step 454.

[0161] A block diagram is shown for the frame of the video which is on a monitor like the video display monitor 61 shown in drawing 1 A and is displayed in drawing 21. The frame of the video of drawing 21 is drawn noting that there are the tree 465 the person 467 the bird 468 and the two clouds 471.

[0162] In drawing 21 the field in a frame is limited including the head of the bird 468 and the person 467 including the field 466 which encloses the tree 465 the field 469 which encloses people and the field 470 and the overlapped range 469. These fields were drawn using the graphic pointing device 33 shown in drawing 1 A. +4 of the field 470 containing the head of +2 and bird and person of the 466 field-5 containing a tree and the field 469 including people may be contained in the quality priority which the user set up by drawing 21. An user definition priority is not specified therefore the priority un"given a definition" is specified as the other field

of the frame of the video in drawing 21. It is the purpose adjusted so that it may become equal to the field un"given a definition" to the number of bits which produces the bit total of encoding of the frame as which an user definition priority is specified as a result of encoding of the origin of a frame so that it may mention laterAfter changing the quantization level of an user definition fieldit is a field corrected first. When a user does not wish change of the quality of the field of a framethese fields will be marked if priority zero are set up by the user.

[0163]In the example of this inventionif a priority field is limited within the frame of videoa user will use a graphic pointing device at the beginningand will create a rectangular area on the frame as which video was displayed. The field defined later is stored on the field defined previouslyand may overlap with the field defined previously. Even if it is explaining that this invention limits a rectangular areait cannot be overemphasized that it is applicable also to the shape which the user containing the field which includes instruction of this invention for a curve like a circle or an ellipsethe field of a polygon like an octagon or a hexagonthe other curvea straight lineor its both defines. A user defines the priority of a fieldafter defining each field. Insteadthe user can define the shape of all the fields and can specify a priority as these fields after that.

[0164]When a user defines a priority in a field in early stagesa field corresponds to pixel positions. Howeverthe process of assigning the priority shown below operates by a macro block unitand a macro block is a unit used for digital video encoding. Thereforethe priority of a field is assigned to the macro block when a macro block is in an user definition field and on a periphery. The person skilled in the art of a Prior art can ask for the macro block corresponding to an user definition field using simple mathematics. In the case of the periphery of an user definition fieldwhen a side part intersects a macro block outside a fieldas an alternative plan of only putting in a macro block in an user definition fieldWhen not less than 50% of macro blocks have the still more precise process of judging whether a macro block being put in in a fieldin an user definition fieldWhen a macro block is put in in an user definition field and 50% or less of macro blocks are in an user definition fieldit can perform by eliminating a macro block from an user definition field.

[0165]After a field and its priority are definedthe user can change duplication of a fieldwhen you wish. For examplewhile a user presses the function key determined a priori on the field 469 of drawing 21it clicksThe field 469 is changed so that it may come on the field 470the priority of the field 469 is assigned to the whole personand the priority assigned only to the bird 468 to the field 470 is set up. When inputted by the usera stack priority is specified as an overlapped range. This stack priority is adjusted if neededwhen an user definition field is addeddeleted or corrected.

[0166]After a field is definedit can ** a user remaking more than ** of a field according to hopeor moving in a field using the graphic pointing device 33and the field defined by the user can be deleted using the graphic pointing device 33. It should annotate to only the user definition priority being deleted and needless to say the original video data in the field under deletion not being deletedeven if the

field is deleted.

[0167] So that it may explain to said paragraph about what quality is changed for in a period between automatic encodings or after automatic encoding The user can review encoding video an input is inputted [whenever interested frame or period occurs a user only presses a key by a keyboard or] from another device and it is shown that a specific frame or period is important. Later a user returns in the frame or period can spend time further can examine a period and can change the feature within the period according to hope.

[0168] The process by which a user defines the field in the frame of video was mentioned above about the simple case only with one frame of the data corrected by the user. However since it is very monotonous and tedious that one frame works the frame of the data in a video sequence at a time at once if this invention is used a user defines a field by the first period and can define a field corresponding in a next period. On all the frames of the middle video of the first frame and the last frame after it. The field corresponding to the field defined as the first frame and the last frame by the user is created. An interpolation process is performed so that the same priority as the first field the last field or the field of the both or the priority based on the priority of the first frame and the last frame may be specified. The stack priority of an overlapped range is maintained between interpolation processes. After the field is automatically created by the intermediate frame the user can change by changing the shape of a priority and a field according to hope can add a new field or can delete a field.

[0169] Below the example of how the interpolation process operates is shown. A user is made to choose the priority field of two frames in the time instance t_1 and t_n . After a user assigns the field corresponding to t_n for a field to the frame t_1 the coordinates of these two fields are interpolated on an intermediate frame in order to gain the field specification of all the frames between t_1 and t_n . It not only can perform interpolation to the actual field under definition but it can perform interpolation to a quality priority. A function is used for an interpolation process temporarily [simple]. A user defines the field A of the frame t_1 with the point a_1 there is point a_n corresponding to the field A to which frame t_n corresponds and it is the intermediate frame a_2 and a_3 ... Level vertical axis defined as the corresponding point in $a_{(n-1)}$ as follows

And the following vertical vertical axes

It is assumed that it is *****. In this case h and v correspond to the level vertical axis in question and vertical vertical axis of a point. In this process only the vertex of the polygon which limits a field is used and a vertex is interpolated. And a field is limited by the peak. Other interpolation techniques other than linear interpolation can be performed.

[0170] The required number of times can perform said interpolation process until all the fields of all the frames in the segment of an editing object are defined appropriately. Furthermore a user moves in the field determined with interpolation

of the middle processor this process. The size of a field can be changed and an interpolation process becomes recursive by the fine adjustment process generated in the first frame and the changed middle process so that it can perform once again in the process of middle frame and the last from it. An interpolation process is used in order to mainly define the field which crosses a frame and where it moves. However defining once the field which is standing it still and is used within the sequence of a frame is also included in this invention.

[0171] After a field is defined and an user definition priority is inputted based on the priority inputted by the user and the old quantized value of a macro block the new quantized value of each macro block is estimated at Step 456 of drawing 20 or is calculated. This is determined according to the relation explained to drawing 22 in which the fraction of the quantized value of the origin used by each of an user choice priority is shown. For example if a user inputs priority zero the fraction produced as a result of the quantized value of the origin used is 1 and it means that there is no change in the original quantized value. If a user chooses priority -3 in order that the original quantized value of the macro block to which the fraction of the original quantized value is set to 1.58 and the priority 3 is set may calculate the quantized value produced as a result it means that multiplication is carried out by 1.58. Since priority -3 means deterioration of quality in order that a quantization level may attain deterioration of quality a quantization level must be pulled up that there must be still few bits used for a macro block. If a positive priority like 3 is inputted by the user on the contrary the multiplication of the quantized value of the macro block from the original formation of an automatic item will be carried out by 0.42 which reduces a quantization level. It means that much more bits are needed for expressing the image in a macro block therefore image quality of the quantization level reduced improves. (-5100)(-41.75) the point on the line shown in drawing 22 It is ((-31.58)(-2 and 1.42)(-1 and 1.25)(0 and 1)1-0.72)(2 and 0.58)(3 and 0.42)(4 and 0.25) and (5 and 0.01). Mapping explained to drawing 22 is based on the linear relation of this direction of [between q-level and a priority]. The point called for by drawing 22 is called for through an experiment and various fractions can be used with an user choice priority according to hope.

[0172] Priority -5 and 5 aim at acquiring the minimum possible quality and the highest possible quality respectively. The minimum quality is a result of 62 in case the maximum permissible dose child-sized level 112 and a linear quantization scale in case a nonlinear quantization scale is used are used for MPEG-2 encoding. The minimum quantization level used in order to attain the top quality is 1.

Therefore although the value which is separated from a actual permission q scale from the relation explained to drawing 22 may arise these calculated q values are only stored on the saturation level. For example the quantizing scale 20 should be set as a macro block and please assume at the beginning that it chooses so that a user may be set to -5 in the priority of a macro block. In order that the quantized value 20 by which multiplication is carried out by 100 may produce the new quantization level 2000 which is separated from a quantizing scale a peak child-sized value or a saturation content child-sized value is used.

[0173] In addition to using a saturation level as mentioned above by this invention the preventive measures of the I frame of MPEG encoding frames and the B frame are taken. This invention restricts the highest quantization level of I type frame so that it may be set to 96 in the case of a nonlinear quantization case and it may be set to 58 at the time of linear quantization scale use. Similarly in the case of a nonlinear quantization scale in the case of 112 and a linear quantization scale the highest quantization level of a type B frame and a P type frame is 62. A linear quantization scale and a nonlinear quantization scale are defined by the MPEG standard. I frame type a p-frames type and the minimum quantization level used for B frame type each are 1. Needless to say the fraction of the user definition input shown in drawing 22 with a block diagram and the original quantized value is correctable in order to satisfy the requirements for a user.

[0174] After a new quantized value calculates for every macro block at Step 456 it asks for the number of bits for every frame produced at Step 458 as a result of the quantity Mr.-ized machine value calculated at Step 456. Although the quality of other fields deteriorated in the example of this invention this step is important in order to have to gain the same number of bits as each frame after the quality of some fields improves. In order to ask for the number of bits produced as a result of a new quantized value it is necessary to set up the function which estimates appropriately the quantization level to the number of bits produced as a result of every macro block. In order to search for the related positive one between the number of bits and the quantization level which are produced as a result examination of the experiential result of a video encoder like MPEG-2 encoder is carried out and the called-for relation is based on the actual experiential result. The simulation result of various quantization levels which use many video sequences and a frame type was carried out. Based on these results the following relation nature between the numbers of bits (b) used when encoding the quantization level (q) and macro block of a macro block was judged to be as follows according to the frame type.

[0175]

In this case in the case of I type frame in the case of B type frame and P type frame it is $m = 0.1$ in $m = -0.75$. In the case of the I frame frames and the B frame it is $n = 15$ and 14.5 and 13.8 respectively. These values are equalized by many video sequences so that it may be annotated above. Said equation works as an approximation standard which excelled for predicting operation of an encoder and the relation between the number of bits and a quantization level is linearity in a log area. Needless to say if exact the other relation between a quantization level and the number of bits produced as a result can also be used when asking for the number of bits produced from the specified quantization level including the rate quantization function of others which are written in this specification.

[0176] Said process explains how to ask for the number of bits produced from the specified specified-amount child-ized level. However when two or more encodings

are used the information from new encoding can be used and a data rate can be made still more exact to a quantization function. There is a 1st video encoding path used for collecting statistics of the video which should be encoded between initial automatic encoding processes, and another actual point of the data rate to a quantization function sets between the 2nd path that is actual encoding at b to q plot corresponding to said equation which describes number-of-bits b created from specified q -level — it is gained like. After data is encoded according to user specification selection so that it may explain below according to whether an improvement of the quality expected of it is filled the video produced as a result is received by the user or is refused. The q -level-related actual data to the bit estimated whenever it re-encoded is created. When new q -level quota q' is specified the called-for data rate changes still more correctly about this point based on the additional data obtained by re-encoding. When new quantization level and q' corresponds between two points obtained before in order to acquire the estimate of the number of bits outputted on a new q -level linear interpolation can be performed in a log area. When the new quantization level " q " does not correspond between q -levels determined as beforehand [two] a model can be used for estimating the number of bits which extrapolates to new quantization level q' and is created from nearest q -level at the time of encoding. It comments on the probability of acquiring the estimate more outstanding to the outputted bit from an encoder becoming high so that the time when data is re-encoded is long.

[0177] After the number of bits of each frame produced from the estimated quantized value is called for at Step 458 of drawing 20 it judges whether a bit estimated number is close to the original number of bits of a frame enough at Step 460 and if required the number of bits is corrected. This is performed according to the process shown with a block diagram by drawing 23–25C.

[0178] If Step 460 is performed Step 472 of drawing 23 will ask the beginning for the difference between a bit estimated number and the number of bits of the original encoding. This is calculated according to the following equations.

[0179]

In this case the rough number of bits of the frame by which B' was corrected the number of bits of the frame which produces B from the original encoding as a result and D are differences. It cannot be overemphasized that the value of B other than the number of bits produced from the original encoding as a result can be used. For example the user can set B as the number which is the quantity determined as beforehand [larger] than the number of bits produced from the original encoding as a result when wishing for the increase in the total bit number of a frame. Similarly the user can also set up B it be less than the original number of bits of a frame. Therefore the recorded frame may consume many bits or few bits from the original encoding according to the possibility of restrictions of a storage area the minimum quality restrictions and a video buffer underflow. When the quantized value of the field in a frame is not changed only the number of bits

produced from the field which had the quantized value changed as a result can be analyzed and the field which was not changed can be disregarded. And this difference is analyzed at Step 474 and if a bit is judged [many / too] a flow will progress to B for the process explained to drawing 24 A–24C to pull up the quantization level of a macro block in order to reduce the number of bits. If the bit in the frame corrected at Step 474 is judged to be too few the quantization level of the macro block of a new frame needs to be reduced in order to create much more bits according to the process F shown in drawing 25 A–drawing 25 C with a block diagram. When it is in the threshold level which can permit a difference when other it is not necessary to correct the quantization level of a macro block and a process returns in order to perform Step 462 of drawing 20. When keeping the number of bits in a frame the same to the re-encoded video it is possible that a bit is shifted from a certain field to another field.

[0180] The difference D is analyzed by comparing a difference with the threshold level T in Step 474. When the number of bits which produces threshold level as a result of change of the quality of a field is less than 0.001 of a total bit number a difference can be permitted and a process is set as $0.001 \times B$ which means that it can stop. The stop threshold level which shows that a difference is permissible needless to say is called for with an option it can pull up according to hope or can be reduced or can be based on both the original number of bits of a frame and the number of bits produced from an user choice priority as a result. The process explained to drawing 24 A–drawing 24 C and drawing 25 A–drawing 25 C Except for the point of being used in order that drawing 24 A–drawing 24 B may reduce the number of bits produced as a result of user choice quality and being used in order to make the number of bits which drawing 25 A–drawing 25 C produces as a result of user choice quality increasing it is extremely similar. The general outline of the process performed by drawing 24 A–drawing 25 C is described here before explaining the details of these processes. At the beginning a label is attached to the macro block in the frame of video as one of four kinds "change is [1] improper". ["no giving a definition" negative being a "priority" positive being a "priority" and] When the priority by a user is not specified as a field or when a user can assign the status un"given a definition" by a system it is considered that the field is a type un"given a definition." If a user is between -5 and -1 and assigns the priority containing -5 and -1 it will be thought that the macro block in these fields is "a negative priority." If it is between 1 and 5 and the priority containing 1 and 5 is assigned to a field it will be thought that "the positive priority" is set to those fields. When all the fields assigned by the priority 0 or systems can finally show that a user must not make the quality of a field change by other methods of a certain it is considered that the macro block of those fields is a "change is impossible" type. A label is attached to each of the macro block in the inside of each of four kinds of macro blocks by the index i so that it may have a set of an index original with each of four kinds of inside. The index i of zero is assigned to the first macro block and the index n-1 is assigned to the last macro block. In this case N is a total of the macro block within a certain field kind. For example the

index of 0 to n-1 is attached to each of the macro block in these fields to all the fields un"given a definition." When there are three or more fields un"given a definition"it annotates to the two different indexes i= 0 not being found in the field un"given a definition." The number of the macro blocks to which the index 1 attached only one macro block to which the index 0 was attached to the field un[all the / "not given a definition"] is one.

[0181]When there are too many bits in a frame after a user determines the priority of a fieldthe quantization level of a frame must be pulled up in order to reduce the number of bits of a frame. The original quantized value of a macro block is deleted from the log file of a macro block layer with the format shown with a block diagram in either drawing 10 (A) or drawing 10 (B). Insteadan encoder is rerunand a quantized value is calculated when the quantized value is not memorized after execution of the origin of an encoder. Howeverin another alternative planthe encoded bit stream is decrypted and the quantized value of the original encoding is calculated. In order to reduce the number of bitsthe procedure operates by pulling up one macro block of quantization levels for every macro block of the type un"given a definition" at a time at once 1 quantization level. When a problem is not solved nowq level of each macro block un"given a definition" increases at a time again one time by one macro block at once. The peak child-ized level determined a priori is set upand it can increase exceeding the quantization level of a non-macro block. If all the macro blocks of the type un"given a definition" reach this maximum levelit will be similarly increased by the field of "a negative priority" and "a positive priority" one time. firstwhen the problem of bit difference is not solved for all negative priority macro blocks by 1 quantization-level raising ** and thisthe quantization level of the macro block of a "positive priority" macro block is increased one time. When a problem is not solved nowit is again increased by the quantization level of a "negative priority" macro block one timeand when a problem is not solved nowthe quantization level of a "positive priority" type macro block is increased one time. This process continues until the quantization level of all the a "negative priority" type and "positive priority" type macro blocks can pull up to said maximum level. when a bit difference problem is not solved nowuntil a bit difference problem is solved for the quantization level of a "change is impossible" type macro block -- at once -- 1 quantization-level raising ****. In all probabilitya bit difference problem will be corrected before a "change is impossible" type quantization level begins to be changed. Since there is too little number of bits of the frame to which an user choice priority is set and it increases the number of bits according to the process of drawing 25 A-drawing 25 Cwhen the quantization level of a macro block needs to be reduceda similar process occurs.

[0182]If the flow chart shown in drawing 24 A performed by drawing 23 with a block diagram is shortly referred to when there are too many produced bitsa process will be started by drawing 24 A by Step 480 which sets a block type as the type un"given a definition." And it is set as the start points which mean that the index i is set up equally to zero in a macro block index at Step 482. And the

set-up macro block type (for the first time with a flow chart) Step 486 which judges whether the maximum prior definition quantization level is set as the quantization level of the macro block (index = macro block in i) of the problem of the 1st macro block of the type un"given a definition" is performed. In the quality of the macro block corrected the peak child-ized level of this correction process is too expensive and must not be set up so that there may be no big degradation. If possible what it is going to degrade quality for in order to decrease the uniform number of bits is desirable so that a big difference may not appear in the quality of a macro block. Therefore this maximum is set as 112 when a linear quantization scale is used and a nonlinear quantization scale is used for the quantization level 62 by MPEG-2 encoding etc. However the other value can also be used for a peak child or a vessel value. If the quantization level of the macro block in question is judged not to be the maximum at Step 486 a quantization-level increment of the quantization level of the macro block in question will be carried out at Step 488. And the bit estimate of a frame is recalculated at Step 490 using the aforementioned bit rate quantization function. And in Step 490 the elephant sentence of the index of a macro block is carried out. When the index of a macro block shows that the macro block of the last of a macro block type is not processed it acts to Step 484 as loopback of the flow. When other from Step 492 a flow progresses to Step 498 which resets a block index to zero and shows the 1st macro block.

[0183] If a flow returns to Step 484 the bit estimate recalculated at Step 490 will be used in order to recalculate the aforementioned difference D compared with threshold level so that it may mention above. When bit estimate is not too high the number of bits of the frame which has an user choice priority specified is corrected and a flow returns to a calling process. When other a flow progresses to Step 486 which judges once again whether the quantized value of the macro block (the index 1 is set up) in question is the maximum.

[0184] If it judges that the quantized value of the macro block in question is the maximum at Step 486 it is not necessary to pull up the quantization level of the block and the increment of the 1-block index i will be carried out at Step 494. If it judges that the index of a macro block type has not passed the last block at Step 496 it will be judged whether the peak child-ized value is set as the block to which the index by which increment was carried out by a flow progressing to Step 486 was attached. If it judges that a macro block is a macro block of the last of the macro block type in question at Step 496 when other Step 498 will be performed and a block index will be reset by the 1st macro block (i= 0). And a flow progresses to the process C explained to drawing 24 B.

[0185] At drawing 24 Bit is investigated at Step 500 whether a block type is a type un"given a definition." When it is a type un"given a definition" it is judged whether a maximum of q levels are set as the macro block un[all the / "not given a definition"] at Step 502. Since it is impossible to perform adjustment further to the macro block of the type un"given a definition" when a maximum of q levels are set to all the macro blocks un"given a definition" A macro block type is changed

into "a negative priority" at Step 504a flow returns to the process D shown in drawing 24 A with a block diagram and Step 484 is performed. If it judges that a maximum of q levels are set as no macro blocks of the type un"given a definition" at Step 502 when othera flow will return to a process without changing the type of the macro block under correction. This procedure is set as the peak child-sized level which produces the block type with which it continues until the number produced as a result of a frame becomes threshold level within the limits or all the macro blocks of the mold un"given a definition" are changed into "a negative priority" as a result.

[0186]It is judged whether Step 506 is [a block type] "a negative priority" as Step 500 "defined" the block type and judging. As for Step 506 in the case of "a negative priority" a block type judges whether a maximum of q levels are set as all the blocks of "a positive priority." When the result of judgment of Step 508 is negative Step 510 is performed a block type is set as "a positive priority" and a flow progresses to the process D shown in drawing 24 A with a block diagram.

[0187]If it judges that a maximum of q levels are set as all the macro blocks of "a positive priority" at Step 508 it will be investigated whether a maximum of q levels are set as all the "negative priority" type blocks at Step 512. When not set up a block type is not changed but a flow returns to the process D of drawing 24 A. To all the "negative priority" type macro blocks at Step 512. When it is judged that a maximum of q levels are set up a peak child-sized level is set as the macro block of "none of all the giving a definition" a negative priority and "a positive priority" and a block type is set as "change is improper" at Step 512. And Flo progresses to the process D of drawing 24 A.

[0188]After Step 506 of drawing 24 B is judged to be negative the process E and Step 520 of drawing 24 C are performed. It is judged at Step 520 whether a block type is "a positive priority." When that is right it is judged whether a maximum of Q levels are set as all the blocks of "a negative priority" at Step 522. When not set up a block type is set as "a negative priority" and a flow returns to the process D shown in drawing 24 A with a block diagram. If it judges that a maximum of q levels are set as all the "negative priority" type blocks at Step 522 it will be investigated whether a maximum of q levels are set as all the "positive priority" type blocks at Step 526. When not set up a flow returns to the process D shown in drawing 24 A with a block diagram. When othera peak child-sized level is set as all the "negative priorities" and "positive priorities" with a block un[all the / "not given a definition"] a block type is set as "change is improper" at Step 526 and a flow returns to the process D shown in drawing 24 A with a block diagram.

[0189]If it judges that a block type is not "a positive priority" at Step 520 so the block type must be "impossible [change]" and it will be judged whether a maximum of q levels are set as all the "change is impossible" type blocks at Step 530. Since all the block types were set as the peak child-sized level an error arises without solving the problem that there are too many bits in the frame to which an user choice priority is set when set up. If all the blocks are set as a maximum prior definition q level the number of bits which does not exceed the

original encoding of a frame should arise. When judgment at Step 530 is negative a flow returns to the process D explained to drawing 24 A.

[0190] If the process of drawing 23 judges that there is too little number of bits in a frame with an user choice priority field the process of drawing 25 A-drawing 25 C will be performed. Drawing 25 A-drawing 25 C is reduced instead of the ability to pull up in order for a quantization level to make the number of bits of a frame increase. When the quantization level of a macro block is reduced it is based on the process shown in drawing 24 A-drawing 24 C with a block diagram except for the point that the minimum level determined as beforehand [of a quantization R level] like 1 cannot be exceeded. Except for the point of difference on which it commented above since the process of drawing 25 A-drawing 25 C is the same in the process of drawing 24 A-drawing 24 C further explanation of drawing 25 A-drawing 25 C is not given in order to expect ****.

[0191] The correction process described by drawing 23-drawing 25 C is one form with which this invention operates. However the substitute means which adjusts the quantization level of the macro block of a different kind is also possible. For example it can be uniformly increased by the macro block a "negative priority" type a "positive priority" type and "change is impossible" type instead of changing simultaneously the quantization level of the macro block of "a negative priority" and "a positive priority" as mentioned above. After being able to pull up instead until a "negative priority" type quantization level reaches a peak child level when there are too many bits the quantization level of a "positive priority" type macro block can pull up. Since the number of bits is increased before the level of a "negative priority" type macro block is pulled down when there is too little number of bits to which an user choice priority is set on the contrary the quantization level of a "positive priority" type macro block may be reduced. In the two latter cases change of the macro block of the type un"given a definition" may be generated in before "a negative priority" or a "positive priority" type macro block between or the back.

[0192] After the process of drawing 23 A-25C is performed a flow returns to Step 462 of drawing 20 which re-encodes video using a new quantized value. It is the purpose of bringing the number of bits produced as a result of a re-encoding frame close to the number of bits of encoding of the origin of a frame as much as possible this re-encoding uses a rate control function and while tuning a quantized value finely it is executed. A rate control device process is explained to the following another section in detail. During re-encodings since encoding artefact may occur a special means can be performed in order to reduce such encoding artefacts so that it may explain to the following another section.

[0193] The process of space manual edit is generated by workstation 30 of drawing 1 A of the example of this invention. However when space manual edit does not occur simultaneously with a video encoding process the workstation 10 can carry out the matter of the time manual edit without degrading the rate of a video encoding process. It should annotate to this invention being described in detail as changing the image quality of various fields in a frame. It cannot be

overemphasized that the concept of this specification is expandable to shifting the bit in the field of a certain frame to the field of another frame. Space manual edit is applicable together with the aforementioned time manual edit.

[0194]Space manual edit was explained as keeping the same as the original encoding of a frame the number of bits of the frame to which the user appointed field priority is set. However when there is sufficient storage area capacity there are more edit segments re-encoded when he wished than the original number of bits of a frame or they can consume the number of bits determined as beforehand [little]. In these cases it is desirable to understand the number of bits which can change the changed frame therefore the process shown in drawing 24 A-drawing 25 C with a block diagram will operate until the set-up difference is attained. Therefore the arbitrary bit totals of re-encoding of a frame consume the arbitrary quantity which there is mostly between 0% containing both few [that it is 1% as more as 5% or] bits and 20% and exceeds the original encoding or than which it is less. [little]

[0195]Explanation of space manual edit was given about change of the video after video was encoded. However the instruction written in this specification is applicable also to a system without an initial encoding process.

[0196]In order for space manual edit to attain the same number of bits as the frame acquired before after assigning a quantized value based on an user choice priority it operates by adjusting the quantization level of the macro block of a frame if needed. By contrast the time manual edit described by the aforementioned paragraph deletes the common pool of a bit from the section of video and redistributes those bits from a common pool according to an user choice priority. All the instruction about the process in which space manual edit is performed about a certain frame is applicable to the concept of changing the quality of the frame of video by time and it is also the same as when reverse so that it may be explained to the paragraph of time manual edit. For example in order that space edit may attain change of quality it has been explained as correcting a quantization level but the number of bits is adjusted by the paragraph of time edit. Since it depends for the number of bits of a frame or a macro block on a quantization level directly correction of the number of bits and a quantization level attains the same result therefore time edit and the space edit can correct the number of bits a quantization level or its both.

VIII. Re-encoding A. of the video after quality change in order to control the bit rate of the video encoding device 50 shown in the bit quota control diagram 1A when quality is changed within an independent frame at a block diagram A control device like all the special hardwares that can perform the workstation 10 or the same control facility loads the rough number of bits or the estimated number of bits for every slice of HE in a quantization level i.e. q -level and a rate controlling expression encoder. And an encoder starts encoding and a quantization level is adjusted after measuring the accuracy of the created number of bits as compared with the estimated number of bits called for before. The estimated number of bits is called for using a rate quantization function so that it may be mentioned above

by the paragraph about space manual edit and time manual edit. This process is continuously performed to a set of n one-time macro block each set is indicated to be $MB(i)$ and it is as follows.

[0197]

$1 \leq i \leq (\text{total number}/n \text{ of a macro block})$ Since minor superfluous anticipation of the number of bits created can absorb using bit stuffing when re-encoding a set of the last of (27) macro blocks A rate control device guarantees that the total of the bit created in the re-encoding process becomes shorter than the length encoded first.

[0198] Drawing 26 is a flow chart explaining the method for re-encoding a set of a digital video with the quantization level assigned a priori. Adjustment is added to the quantization level to which the quantization model was however assigned a priori in the number of bits created by re-encoding based on that under anticipation to the actual number of the created bits correctly. In drawing 26a rate controlling expression coding process specifies the 1st set of n macro block as a set of the present of a macro block by setting up $i = 1$ at Step 702 after starting at Step 700. In Step 704 a perfect set of the number of bit anticipation used by re-encoding of each set $MB(i)$ of a macro block and $EB(i)$ are loaded. And $MB(i)$ is re-encoded at Step 706 and the number of the actually created bits is memorized as $GB(i)$. The percentage difference between created bit $GB(i)$ and anticipation bit $EB(i)$ is calculated as an estimated rate and $PB(i) = GB(i)/EB(i)$ at Step 708. The system can calculate number $EB'(i)$ accumulation creation number-of-bits and $GB'(i)$ of anticipation of an accumulation bit to the set by which the macro block was re-encoded altogether. In this case $EB'(1) = EB(1)$ $GB'(1) = GB(1)$ $EB'(i) = EB'(i-1) + EB(i)$ and in the case of $i > 2$ it is $GB'(1) = GB'(i-1) + GB(i)$. Therefore $PB(i)$ is calculable as an alternative plan of explained Step 708 by using $PB(i)/EB'(i)$ using number $EB'(i)$ and accumulation creation bit several GB' of accumulation bit anticipation (i). This serves as a more exact display of how the re-encoding process as the whole is filling the number of target bits instead of it being shown whether the specific set with a macro block was encoded correctly however. In Step 710 an approximate rate is used and the set next to a macro block and the correction coefficient of q -level in $MB(i+1)$ are judged.

[0199] This correction coefficient is calculated using one of two kinds of methods. The 1st correction coefficient deciding method follows the list of rates in the table of the rate value comparatively called a look-up table rearranged [simple] and judges comparatively whether which two an approximate rate corresponds between 0. The index j of the one where two numbers in a table are larger is used as a correction coefficient index. A correction coefficient index is used in order to look for correction coefficient $** (i)$ within a correction coefficient table by choosing the j -th element in a table as a correction coefficient and the set next to a macro block is adjusted by $** (i)$. the rearranged following values are comparatively specified as a look-up table and RT for example -- $RT = \{0.850.900.951.01.051.101.151.2 \text{ and } 1.5\}$

The following values are specified as correction coefficient table CT .

[0200]

CT={-3-2-1012349}

Drawing 30 is a flow chart explaining the process of asking for a correction coefficient index from a look-up table comparatively using an approximate rate. Only for a block diagram it is assumed that it is $PB(i)=0.87$. The index j is set up at Step 722 at the beginning. The process in comparison with $RT(j)$ starts $PB(i)$ and $PB(i)=0.87$ is compared with $RT(1)=0.85$ by Step 724 so that it may be defined as said look-up table. When 0.87 exceeds 0.85 a process follows Step 726. Since it is judged that there is the remaining index at Step 726 control is passed to Step 728 and 1 increment of the j is carried out. Since $PB(i)=0.87$ is $RT(2)=0.90$ now when it returns to Step 724 control is passed to Step 730 which ends search of a correction coefficient index.

[0201] Drawing 31 is a flow chart explaining how to discover a value within the look-up table as which the index of TEBURUHE is specified. In drawing 31 in order to calculate a correction coefficient in accordance with the 1st method the index j which is 2 is used and the corresponding entry in CT is discovered. If it reads in the 2nd position in CT it will be judged that a correction coefficient is Step 736-2. Therefore q -level of MB ($i+1$) is changed by -2. Similarly since the index 7 corresponds to the input of the beginning in larger RT than $PB(i)$ in the case of $PB(i)=1.12$ the correction coefficient of CT in the 7th position is 3. Therefore q -level of MB ($i+1$) is changed by adding 3 to MB ($i+1$) for example by adding 3.

[0202] In the 2nd correction coefficient deciding method when the percentage of a set of the re-encoded macro block calculates a correction coefficient a process similar to the 1st method is used except for the point used as the 2nd parameter. A correction coefficient is read in a matrix instead of reading a correction coefficient in a table. More gradual correction is added by the beginning of the sequence of a frame by using the percentage of a set of the re-encoded macro block or the percentage of a set of a macro block which is not re-encoded and if required more remarkable correction can be added at the last. By this the system can correct now more correctly change between a quantization model and the actually created number of bits. The inversion relation between the number of sets of the macro block which remains without being re-encoded and the meaning of correction is used and unnecessary correction will be avoided if superfluous anticipation within the set with macro part AROKKU assumes that each other is offset by the too little anticipation by another set of a macro block.

[0203] As an example of correction coefficient matrix coal tar mixture of the 2nd correction coefficient method it is assumed that coal tar mixture is defined as follows.

[Equation 1]

[0204] By the beginning of the sequence of a frame when a rough ratio shows superfluous evaluation of the required number of bits it is necessary to reassign no intact bits immediately to the next slice. In the case of $k_{i+1}MB(k+1)$ will be

compulsorily pulled up to an unnecessary still higher quantization level by MB (k) which consumes the bit beyond anticipation when an intact bit is immediately consumed by MB (i+1). An intact bit is ahead carried from MB(i) in a re-encoding process until it is needed and is used by MB (k).

[0205] Drawing 32 is a flow chart explaining how a correction coefficient is calculated using both a rough ratio and the number of sets of a macro block which remains and which are not re-encoded. When it is judged before that it is equal to 2 when the case j of superfluous evaluation of the needed number of bits i.e. a correction coefficient is $PB(i)=0.87$ superfluous evaluation judges [of the beginning of a re-encoding process] whether par SENTODE generating is carried out at Step 744 of drawing 32. The element of (21) of coal tar mixture is -1 therefore is chosen as a correction coefficient. thereby by the 1st method the correction effect becomes blunt in early stages of the re-encoding process which came out and had -2. However since 99% of re-encoding processes are performed and the element (977) which is 3 is chosen when a rough ratio is $PB(i)=1.12$ it corrects thoroughly at the last of a re-encoding process.

[0206] If it returns to drawing 27 (A)–drawing 27 (C) the number of macro blocks in a set with a frame to which q-level was assigned a priori will be explained as another example equal to the number of macro blocks in one slice. Estimated number $EB(i)$ of a bit is calculated as mentioned above and after being loaded in a memory used by MB(i) of drawing 27 (A) and a rate controlling process of the macro block i MB(i) is actually re-encoded. The creation number of bits and $GB(i)$ are calculated a rough ratio is calculated and the search for a correction coefficient is taken out. As shown in drawing 27 (B) MB (i+1) is corrected by $** (i)$ and a process is repeated for MB (i+1). MB (i+1) is estimated and calculated by drawing 27 (C) and change of MB (i+2) after the rough ratio was calculated is shown in it. This process continues until all the macro blocks and partial macro blocks are re-encoded and are corrected.

[0207] Another example is shown with reference to drawing 27 (A) and drawing 27 (D). In this example a macro block number per set is the number of macro blocks of relations of two slices. After calculating $EB(i)$ $GB(i)$ and $PB(i)$ a macro block of MB (i+1) which constitutes two slices is corrected by $** (i)$ so that it may be explained to drawing 27 (D). This re-encoding and a correction process continue in a former example until all the macro blocks are re-encoded.

[0208] Change produced by a re-encoding process is explained to drawing 28 (A) and drawing 28 (B). The field A and the field B which are fields of two equal area of the frame N are encoded using a beginning almost equal number of bits. Quality of the field A deteriorates and a user specifies that quality of the field B must improve. However it is desirable to make it the number of bits used for the original encoding become as close to the number of bits of re-encoding as possible. After re-encoding the frame N the same number of bytes created from the original encoding is stored in the frame N but the field A is coded in a smaller number of bits and the field B uses an additional bit. Thereby as a user required quality of the field A deteriorates and quality of the field B improves. In a rate controlling

process of this invention the number of bits by which a frame was re-encoded is brought as much as possible close to the original number of bits by correcting a minor error in an estimated number called for using a rate quantization function. [0209] By drawing 29 by reassigning a bit from two or more frames and giving them to two or more 2nd frames explains change added to a digital video stream. Even if each frame surface product changes a gross area of a set of a frame must still be the same. With a rate control device even if a rate quantization function used in order to ask for a bit estimated number is slightly wrong the number of bits of re-encoding of two or more frames can be finely tuned so that it may become the same as the number of bits from the original encoding.

B. 1. re-encoded while avoiding decryption artefact on an edit point above-mentioned time quality edit operation and space quality edit operation which restore an encoder to the state where it was used when it encoded first. It enables it to change quality of video by re-encoding video with a new quantized value. However decryption artefact which appears as a visible glitch in an edit point of video and which is not accepted may arise only by substituting re-encoded video for original video. Suitable syntax like MPEG-2 syntax is set as a new segment substituted and these visible glitches are generated even if buffer limitations in an edit boundary are filled. A problem is produced as a result of the non-availability of another reference frame (P-frame of a private talk) in the last of the last reference frame (the last P-frame) and an edit segment which comes before an edit segment. Thereby a gap arises between encoding of the 1st set of B-frame in the beginning of an edit segment and the 1st set of an immediately after [an edit segment] B frame and decryption.

[0210] In order to perform edit of digital encoding compression video appropriately it is necessary to determine a suitable position of an encoding bit stream corresponding to an edit point of hope first. the [explaining a utility for this process to access encoded video] -- it was mentioned above by VI paragraph. In order to search for bit offset of the picture in question this is performed in totaling the number of bits of an already encoded frame which is respectively and is used so that it may mention above. A directory can also be used in order to maintain information which shows a strict location of a certain specific frame in an encoded bit stream or a period instead.

[0211] If drawing 33 (A)-drawing 33 (E) is seen shortly the block diagram of the period of video used as an example for explaining operation of this invention is carried out. The original section 0 of encoded video is shown by block diagram in drawing 33 (A). It is expected of replacing 0s 10 m from encoded period 5 m 0s in another video sections (from 5 minutes and a zero second to 10 minutes and a zero second). In drawing 33 (B) the new section of unencoded video which must be substituted for video by which drawing 33 (A) was encoded is shown by block diagram. A dashed line of drawing 33 (B) and drawing 33 (C) is used for expressing unencoded video and a solid line of drawing 33 (A) drawing 33 (D) and drawing 33 (E) is used for encoded video being shown.

[0212] Only a section of the new video shown in drawing 33 (B) is encoded using

standard MPEG-2 in which GOP refers to former GOP. Probably 10 m exists in 0s with 0s 5 m of edit point since an exact reference frame which needs decryption artefact for bidirectional decryption used by MPEG-2 video compression is lacking when substituted for video shown in drawing 33 (A). Although the beginning of GOP is returned and referred to former GOP, this problem exists on both the 5-minute point processed so that it may change with this invention and a 10-minute point for a fact that it progresses to the next GOP and the last of GOP does not refer to it.

DESCRIPTION OF DRAWINGS

[Brief Description of the Drawings]

[Drawing 1 A] The block diagram of the encoder system architecture.

[Drawing 1 B] The figure showing the device for writing in and transmitting the encoded audio picture image data which is created by the device for obtaining the audio and video which should be encoded and an encoding system.

[Drawing 2] The flow chart which shows general operation of the system of drawing 1.

[Drawing 3] The flow chart which shows the process performed in the state of a setup.

[Drawing 4] The figure showing the process performed by a compression state.

[Drawing 5] The figure showing the process performed by an edited state.

[Drawing 6] The figure showing the process performed for a start state a completion state and exit status.

[Drawing 7] The flow chart which shows the process of executing simultaneously the method of determining the number of the audio encoders needed for encoding the audio track of a predetermined number using a multipass video encoder audio encoding and multipass video encoding.

[Drawing 8 A] The figure showing the log file format of a sequence layer.

[Drawing 8 B] The figure showing the log file format of a sequence layer similarly.

[Drawing 8 C] The figure showing the log file format of a sequence layer similarly.

[Drawing 8 D] The figure showing the log file format of a sequence layer similarly.

[Drawing 9 A] The figure showing the log file format of a picture layer.

[Drawing 9 B] The figure showing the log file format of a picture layer similarly.

[Drawing 9 C] The figure showing the log file format of a picture layer similarly.

[Drawing 10] The figure showing the full format of the log file of a macro block layer and the short format of the log file of a macro block layer.

[Drawing 11] The figure showing the process for utilizing the information in a log file.

[Drawing 12] The figure showing the general process performed during the hand regulation of the image quality to a different period.

[Drawing 13] The figure showing the user interface used in order to input the parameter for changing image quality in various periods.

[Drawing 14] The figure showing the general means for calculating the number of

bits corresponding to desired image quality.

[Drawing 15] The figure showing the process used in order to process the section where the greatest or minimum user choice priority is set up.

[Drawing 16 A] The figure showing the flow chart for calculating the number of bits corresponding to the image quality of the hope which is not the greatest or minimum priority.

[Drawing 16 B] The figure showing the flow chart for calculating the number of bits corresponding to the image quality of the hope which is not a priority of the same maximum or the minimum.

[Drawing 17] The figure showing the flow chart of the process used in order to check when an underflow arises as a result of changing image quality.

[Drawing 18] The figure showing the flow chart for judging whether a buffer underflow occurs with the frame after an edit segment for change added within the edit segment.

[Drawing 19] The figure showing the example about the bit rate of the video encoded by the quality which the user chose being changed how and the example about the bit rate of the video encoded by the quality which the user similarly chose being changed how.

[Drawing 20] The figure showing the general process performed between the hand regulations of the image quality of various fields in one frame.

[Drawing 21] The figure showing the example of various fields where the priority which the user chose into one frame of video was set up.

[Drawing 22] The figure showing the graph of the user choice priority over the fraction decimal fraction of quantization of the origin used for the selected priority.

[Drawing 23] The figure showing the process for judging whether whether the difference between the number of bits produced as a result of an user choice priority and the number of bits produced as a result of encoding of the origin of a frame being permissible and a correction procedure must be performed.

[Drawing 24 A] The figure showing the process for correcting a frame when the increment of the quantization level of a macro block is carried out in order to reduce the number of bits since the set-up bit is judged [many / too].

[Drawing 24 B] The figure showing the process for correcting a frame when the increment of the quantization level of a macro block is carried out in order to reduce the number of bits since the bit similarly set up is judged [many / too].

[Drawing 24 C] The figure showing the process for correcting a frame when the increment of the quantization level of a macro block is carried out in order to reduce the number of bits since the bit similarly set up is judged [many / too].

[Drawing 25 A] The figure showing the process for correcting the bit in a frame which are by reducing the quantization level of the macro block in one frame since the number of bits is increased.

[Drawing 25 B] The figure showing the process for correcting the bit in a frame which are by reducing the quantization level of the macro block in one frame since the number of bits is similarly increased.

[Drawing 25 C] The figure showing the process for correcting the bit in a frame

which are by reducing the quantization level of the macro block in one frame since the number of bits is similarly increased.

[Drawing 26] After loading from a memory the number of anticipation of the bit used by each of a set of n macro blocks. In order to count a number for every set of n macro blocks as a result of a bit to calculate the accuracy of the estimate based on the number of anticipation and the number of creation of the bit and to stop in the target size of the sequence of a frame closely. The flow chart which shows the general process of adjusting the quantization level to which the following n macro blocks were assigned beforehand.

[Drawing 27] After one slice is included in each set of a macro block and the 1st set of a macro block is re-encoded (A) the figure showing the picture of the frame which shows the quantization level by which fixed assignment was carried out and (B) After the 2nd set of a macro block is re-encoded the figure showing the picture of the frame of drawing 27 (A) when the 1st correction coefficient is calculated and the 1st correction coefficient is added to the 2nd set of a macro block and (C). The 2nd correction coefficient was calculated and the 2nd correction coefficient was added to the 3rd set of the macro block. As for the figure showing the picture of the frame of drawing 27 (B) and (D) two slices are included in the 1st set of a macro block. The figure showing the picture of the frame of drawing 27 (A) after the 1st set of the macro block was re-encoded when the 1st correction coefficient is calculated and the 1st correction coefficient is added to the 2nd set of a macro block.

[Drawing 28] 1 including a different field which is two by which the quantization level is changed. The figure showing a frame. The figure showing the picture of two encodings of the frame of drawing 28 (A) in which change of the number of bits within the video stream used since a different field which is two by which the quantization level was changed is expressed is shown considered.

[Drawing 29] The figure showing the video stream which shows change of the number of bits used since the president of a university of the re-encoded sequence expresses a series of pictures equal to the original length of a sequence.

[Drawing 30] The figure showing the flow chart explaining how a correction coefficient index is called for from a proportionality look-up table using an approximate rate.

[Drawing 31] The figure showing the flow chart explaining how a correction coefficient is chosen from a correction coefficient table using a correction coefficient index.

[Drawing 32] The figure showing the flow chart explaining how a correction coefficient is calculated using the remaining cardinal numbers by which a rough rate and a macro block are not re-encoded.

[Drawing 33] The figure showing the section of the original video which takes the format by which (A) was encoded. The figure in which (B) shows the new section of the video in the format which must be transposed to the encoded original video and which is not encoded. The figure showing the new section where the video by which (C) has a section where original video is not encoded in the both ends is not

encodedAs for (D)the figure showing the video of drawing 33 (C) in the encoded format which must be transposed to the encoded original videoand (E) are figures with which the replaced video which was encoded is contained in it and in which showing the encoded original video.

[Drawing 34]The figure showing the process of transposing the section of video to the encoding video-data stream used as the video which has the decoding structure which decreased.

[Drawing 35]The figure showing the new section where the video by which (A) has a section where the original video in the falling end is not encoded is not encodedAs for (B)(C) is a figure showing the section of the encoded video which is transposed to the video encoded by the beginning in the block diagram of drawing 33 (A)and a figure showing the video encoded by the beginning of drawing 33 (A) containing the encoded video by which drawing 35 (B) was replaced.

[Drawing 36]The figure which illustrates the process for decreasing decoding structure when editing the video bit stream encoded by encoding the video replaced using a closing group's picture mode.

[Drawing 37]In order that (A) may create the quantization level to a bit rate model in a prior encoding processAs for the figure showing the frame of the unsettled video by which two quantization levels assigned to it are set upand (B)(C) is a figure showing a frame with four quantization levels assigned by a rotation patternand a figure showing a frame with the quantization level of 4 ** distributed in a block formation.

[Drawing 38]The flow chart explaining the process of determining the bit rate to the quantization level which acquired the video signalencoded those signals in advanceand was specified in the prior encoding phase.

[Drawing 39]The figure showing two recorded frames and those corresponding quantization levels to the bit rate.

[Drawing 40]It is a figure showing the outline of the last format of the encoded data which is eventually stored in an optical disc.

[Drawing 41]The figure explaining the disk information file 908 of drawing 40.

[Drawing 42]The figure showing the contents of the data file management information of the data file shown in drawing 40 with a block diagram.

[Drawing 43]The figure showing the details of the data memorized by the data file of drawing 40.

[Drawing 44]The figure showing the data pack in which the data does not need a stuffing packet in itand the figure showing the data pack which needs a packet for padding.

[Drawing 45]The figure showing a reproduction information pack.

[Drawing 46]The figure showing the information included in the data retrieval information 996 on drawing 45.

[Drawing 47]The figure showing the general information of drawing 46.

[Drawing 48]The figure showing the synchronous reproduction information shown in drawing 46 with a block diagram.

[Drawing 49]The figure showing a video pack.

[Drawing 50]The figure explaining the relation between the group (GOP) of a picture and the sequence of a video pack.

[Drawing 51]The figure showing the audio pack encoded according to MPEG audio encoding specification and the figure showing the audio pack built according to AC-3 or linear PCM audio encoding.

[Drawing 52]The figure showing the relation between the audio stream and audio pack which were encoded.

[Drawing 53]The figure showing the structure of a sub picture unit.

[Drawing 54]The figure showing the structure of a sub picture pack.

[Drawing 55]The figure explaining the structure of the sub picture device relevant to a sub picture pack.

[Drawing 56]The figure showing change between the display sub picture unit n and the sub picture unit n+1.

[Description of Notations]

10203040 -- A workstation 21 -- Hard disk 22 -- Digital memory storage 31 -- A console display 50 -- Video encoding device 51 [-- An audio encoder 72 / -- An audio interfacing unit 73 / -- Audio decoder.] -- Playback VTR 52 -- Recording VTR 60 -- Digital capture memory storage 70a-70d

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開2001-128126

(P2001-128126A)

(43) 公開日 平成13年5月11日 (2001.5.11)

(51) Int.Cl. ⁷	識別記号	F I	テーマコード* (参考)
H 0 4 N	5/92	H 0 3 M 7/30	Z
H 0 3 M	7/30	H 0 4 N 5/84	Z
H 0 4 N	5/765	5/92	H
	5/781	5/781	5 1 0 F
	5/84	5/91	N

審査請求 有 請求項の数19 O L (全 80 頁) 最終頁に続く

(21) 出願番号 特願2000-278022(P2000-278022)
(62) 分割の表示 特願平9-505784の分割
(22) 出願日 平成8年4月19日 (1996.4.19)
(31) 優先権主張番号 08/502, 012
(32) 優先日 平成7年7月31日 (1995.7.31)
(33) 優先権主張国 米国 (U S)

(71) 出願人 000003078
株式会社東芝
神奈川県川崎市幸区堀川町72番地
(71) 出願人 500163078
タイム・ワーナー・エンターテインメント・
カンパニー・リミテッド・パートナーシッ
プ
アメリカ合衆国、カリフォルニア州
91522、バーバンク、ワーナー・ブールバ
ード 4000
(74) 代理人 100058479
弁理士 鈴江 武彦 (外3名)

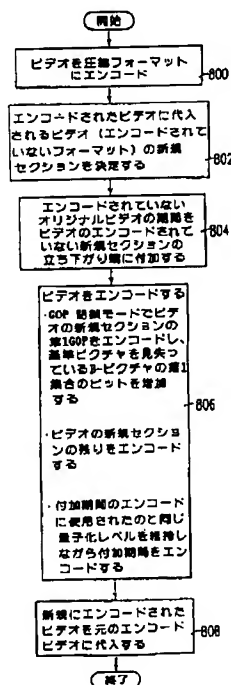
最終頁に続く

(54) 【発明の名称】 エンコードされたビデオビットストリームのセクションを交換するための方法及びシステム

(57) 【要約】

【課題】 編集したときに画質を劣化させないようにする。

【解決手段】 エンコードビデオのセクションを除去するため、エンコードされたビットストリーム内の始点及び終点を決定する。エンコードデータの始点及び終点アドレスは、各ピクチャーで消費されたビット、各ピクチャーのシーケンスヘッダーのビット数、各ピクチャーグループ (GOP) ヘッダーのビット数、及び全スタッフビットを加算するで計算される。挿入されるビデオをエンコードするため、無関係な信号成分のデコードを防ぐ処理が実行される。この処理は編集ポイント前の最終ピクチャーを判断し、編集ポイント後のフレームに関する参照フレームとしてその最終ピクチャーを使用することで行われる。更に最初の量子化値でエンコードされた最初のビデオの短期間が、置き換えビデオの終わりに含まれる。



【特許請求の範囲】

【請求項1】 エンコードされたビデオのビットストリームにビデオを追加する方法であって、前記エンコードされたビデオが修正される始点及び終点を決定する工程と、前記始点及び終点間に挿入するためのビデオ、及び前記終点直後のエンコードされたビデオを示すビデオを含む未エンコードビデオを得る工程と、前記挿入用ビデオの第1グループのピクチャー（GOP）を前記始点及び終点の外側の参照フレームを用いずにエンコードし、第2参照フレームを含まない双方向予測フレームに追加ビットを割り当てる工程と、前記第1GOPより後のGOPを挿入されるビデオの前記終点までエンコードする工程、及び前記終点直後のエンコードされたビデオを示すビデオをエンコードする工程とを実行することによって前記未エンコードビデオをエンコードする工程と、前記エンコードする各工程により作成されるエンコードされたビデオを前記エンコードされたビットストリーム内に配置する工程と、を含む方法。

【請求項2】 請求項1記載の方法であって、一方予測を有するフレームに用いられたビット数を使用して前記追加ビット数を決定する工程を更に含むことを特徴とする方法。

【請求項3】 請求項1記載の方法であって、前記終点直後のエンコードされたビデオを示すビデオをエンコードする工程は、前記未エンコードビデオの最初のエンコード期間中に使用された量子化値に等しい量子化値を用いて、前記エンコードされたビデオを示すビデオをエンコードする工程を含むことを特徴とする方法。

【請求項4】 請求項1記載の方法であって、前記決定する工程は、前記エンコードされたビデオのディレクトリ情報を参照せずに、ビデオフレームに対応するビット数を追加することにより決定する工程を含むことを特徴とする方法。

【請求項5】 請求項1記載の方法であって、カメラを使用してビデオを得る工程と、得たビデオをエンコードして、エンコードされたデータのビットストリームを発生する工程と、を更に含むことを特徴とする方法。

【請求項6】 請求項1の処理により作成されたエンコードされたビットストリームのビデオを格納するためのデータ構造を含むメモリにおいて、

(1) 前記メモリに格納されるデータ構造であって、

(a) 情報の1単位であるメニューセルを複数格納したフィールド、

(b) タイトルメニューの開始セル用フィールド、

(c) オーディオ選択メニューの開始セル用フィールド、

(d) 前記ビデオとは異なるビットマップされた情報を

提供するのに用いられるサブピクチャー情報を選択するための用いられるメニューの開始セル用フィールド、

(e) 前記メモリ内に複数のプログラムが存在するとき、複数のプログラムを選択するために用いる開始セル用フィールド、及び

(f) 前記メモリ内に複数のアングル画面が存在するとき、カメラアングルを選択するために用いるアングルメニュー用開始セルのフィールドを含むデータ構造と、

(2) 表示される各メニューに対するテーブルエントリを含むメニューセル情報テーブルであって、前記前記テーブルエントリは各々、

(a) 前記テーブルエントリに対応するメニューがタイトルメニューか、プログラムメニューか、オーディオメニューか、サブピクチャーメニューか、またはアングルメニューかを示すフィールド、

(b) 選択項目の数を示すフィールド、

(c) 対応するメニューセルの開始部のシステムクロック参照部を少なくとも示すフィールド、

(d) 対応するメニューセルの開始アドレスを示すフィールド、及び

(e) 対応するメニューセルの複数ブロックを示すフィールドを含むメニューセル情報テーブルと、

(3) 各メニューについて表示される情報を格納するためのデータ構造と、

(4) オーディオを格納するためのデータ構造、及び

(5) 請求項1の処理によって作成されたエンコードされたビットストリームのビデオを格納するためのデータ構造、を具備することを特徴とするメモリ。

【請求項7】 請求項6記載のメモリであって、前記メモリは光ディスクであることを特徴とするメモリ。

【請求項8】 請求項1の処理によって作成されたエンコードされたビットストリームのビデオを格納するためのデータ構造を含むメモリであって、

(1) ファイル管理情報を格納するためのデータ構造であり、該データ構造は、(a) 前記メモリ内情報のアドレス及びサイズ情報を格納するためのファイル管理テーブルと、

(b) 情報の1単位である各セルが再生される順番を示す情報を格納するシーケンス情報テーブルであって、該シーケンス情報テーブルは各シーケンスに対するテーブルエントリを含み、各テーブルエントリは、

(b1) 再生された後に終結する完了形式シーケンスか、順番に再生される複数シーケンスの一部である接続形式シーケンスの先頭か、接続形式シーケンスの中心部か、及び接続形式シーケンスの終了部かを格納するフィールドと、

(b2) 前記対応するシーケンス内の複数セルを格納するフィールドと、

(b3) 前記対応するシーケンスの再生時間を格納するフィールドと、

(b4) 前記対応するシーケンスの後に再生されるシーケンスを格納するフィールドとを含むシーケンス情報テーブルと、

(c) セルを示す情報を格納するセル情報テーブルであって、各セルについてのテーブルエントリを含み、各テーブルエントリは、

(c1) 対応するセルの再生時間を示すフィールドと、
(c2) 前記対応するセルのシステムクロック参照の少なくとも一部を示すフィールドと、
(c3) 前記対応するセルのアドレスを示すフィールド、及び

(c4) 前記対応するセルの複数の論理ブロックを示すフィールドとを含むセル情報テーブルと、

(2) オーディオを格納するためのデータ構造、及び

(3) 請求項1の処理により発生されたエンコードされたビットストリームのビデオを格納するためのデータ構造。

【請求項9】 前記メモリは光ディスクであることを特徴とする請求項8記載のメモリ。

【請求項10】 請求項1の処理によって作成されたエンコードされたビットストリームのビデオを格納するためのデータ構造を含むメモリであって、該データ構造はオーディオ及びタイミング情報を含む記述的情報を格納するための構造を含み、

(1) 記述的情報を含むデータ構造であって、

(a) 前記記述的情報データ格納のシステムクロック参照を格納するためのフィールドと、

(b) 他のピクチャーを参照せずにデコードされる対応するビデオピクチャーの開始時間を格納するためのフィールドと、

(c) 前記対応するビデオピクチャーの再生開始時間直前の再生開始時間を有するオーディオ情報のタイムスタンプを格納するためのフィールドと、

(d) 第1ビットは前記オーディオ信号が前記記述的情報データ構造の前に位置するか又は後に位置するかを示し、前記オーディオ情報のアドレスを格納するためのフィールドと、

(e) 前記記述的情報データ構造の前及び後に発生する他の記述的情報データ構造のアドレスを示すフィールドとを含むデータ構造と、

(2) オーディオを格納するためのデータ構造、及び

(3) 前記請求項1の処理によって作成されたエンコードされたビットストリームのビデオを格納するデータ構造と、を具備することを特徴とするメモリ。

【請求項11】 請求項10記載のメモリであって、前記メモリは光ディスクであることを特徴とするメモリ。

【請求項12】 請求項1の処理によって作成されたエンコードされたビットストリームのビデオを格納するためのデータ構造を含むメモリであって、該データ構造はオーディオ及びサブピクチャー情報を格納するための構

造を含み、

(1) サブピクチャーユニットを格納するために前記メモリ内に格納されるデータ構造であって、

(a) サブピクチャーユニットヘッダーを格納するためのフィールドと、

(b) マップされた画像を格納するためのフィールド、及び

(c) 複数のテーブルエントリを含む表示制御テーブルを格納するためのフィールドであって、

(c1) 前記ビットマップされた画像の表示開始時間を格納するサブフィールドと、

(c2) 次のテーブルエントリのアドレスを格納するサブフィールド、及び

(c3) 表示エリア、画素色、画素コントラスト、及びビデオフレームが表示されたときに生じる画素の色とコントラスト変更を含むコマンドグループから選択される少なくとも1つの表示制御コマンドを格納するサブフィールドとを含むデータ構造と、

(2) オーディオを格納するためのデータ構造、及び

(3) 請求項1の処理により作成されたエンコードされたビットストリームのビデオを格納するためのデータ構造と、を具備することを特徴とするメモリ。

【請求項13】 請求項12記載のメモリであって、前記サブピクチャーユニットデータ構造は、前記サブピクチャーユニットデータ構造内の情報を含む複数のパックを具備することを特徴とするメモリ。

【請求項14】 請求項12記載のメモリであって、前記メモリは光ディスクであることを特徴とするメモリ。

【請求項15】 エンコードされたビデオのビットストリームにビデオを追加するシステムであって、

(1) 前記エンコードされたビデオが修正される始点と終点を決定する手段と、

(2) 前記始点と終点の間に挿入するためのビデオを含む未エンコードビデオ及び前記終点直後のエンコードされたビデオを示すビデオを得る手段と、

(3) 前記未エンコードビデオをエンコードする手段であって、このエンコード手段は、

(a) 前記始点と終点の外側の参照フレームを用いることなく、挿入用ビデオの第1グループのピクチャー (GOP) をエンコードし、第2参照フレームを含まない双方向に予測されたフレームに追加ビットを割り付ける手段と、

(b) 前記第1GOP後の複数のGOPを、挿入されるビデオの終点までエンコードする手段、及び

(c) 前記終点直後のエンコードされたビデオを示すビデオをエンコードする手段とを含み、

(4) 前記エンコードする各工程により作成されたエンコードされたビデオを前記エンコードされたビットストリーム内に配置する手段、を具備することを特徴とするシステム。

【請求項 16】 請求項 15 記載のシステムであって、一方向予測を有するフレームに用いたビット数を用いて前記追加ビット数を決定する手段を更に具備することを特徴とするシステム。

【請求項 17】 請求項 15 記載のシステムであって、前記終点直後のエンコードされたビデオを示すビデオをエンコードする手段は、前記未エンコードビデオの最初のエンコード時に用いられる量子化値に等しい量子化値を用いて、前記エンコードされたビデオを示すビデオをエンコードする手段を具備することを特徴とするシステム。

【請求項 18】 請求項 15 記載のシステムであって、前記決定手段は、前記エンコードされたビデオのディレクトリ情報を参照することなく、ビデオフレームに対応するビット数を追加することにより前記始点と終点を決定する手段を含むことを特徴とするシステム。

【請求項 19】 請求項 15 記載のシステムであって、ビデオを得るためのカメラと、得られたビデオをエンコードしてエンコードされたデータのビットストリームを作成する手段を更に具備することを特徴とするシステム。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】 まず、関連出願について述べる。

【0002】 本願は、1995年5月8日に提出された「エンコードされたビデオビットストリームのセクションを交換するための方法およびシステム」と題する米国特許明細書連続番号08/438,014に関連している。本願は又、「オーディオエンコードの数が低減されるオーディオ映像エンコードシステム」と題する米国特許明細書連続番号08/467,991、「既にエンコードされたビデオフレームの品質を手動で変えるための方法およびシステム」と題する明細書連続番号08/466,391、「ユーザが既にエンコードされたビデオフレームの品質を手動で変えるための方法およびシステム」と題する米国特許明細書連続番号08/469,370、「量子化レベルを変化させることにより、ビデオエンコードのビット割当てを制御するレート制御式デジタルビデオ編集法およびシステム」と題する米国特許明細書連続番号08/466,766、および「レート量子化モデルを使用しエンコードするビデオエンコード方法およびシステム」と題する米国特許明細書連続番号08/473,783に関連し、これらの特許明細書は共に1995年6月6日に提出され、本明細書に参考として取り入れられている。

【0003】 本発明はエンコードされたビットストリームセクションが新たなセクションに置き換えられるエンコードシステムに関し、特にエンコードされたビデオセクションがデコード副成分(decoding artifacts)を減少

するように置き換えられるビデオエンコードシステムに関する。

【0004】

【従来の技術】 圧縮法が利用できるようになる前は、オーディオ、ビデオ、及びフィルムは単にオーディオ、ビデオ、及びフィルムのセクションすなわち一部を切り取り、必要に応じてそのセクションを新たな又は代替のセクションに置き換えることにより編集が行われていた。しかし、データをエンコードする圧縮法が用いられるようになると、単純な「カット及び貼り付け」編集は、ビデオフレームを構成するのに用いられる情報として不可能となった。従ってエンコードされたビデオセクションを単純に切り取ることは、変更されていないビデオセクションを劣化させることがあり、新たなビデオがすでにエンコードされたビットストリームに単純に挿入されると、デコード副成分を含むことがある。

【0005】 MPEGビデオエンコードにおいてこの問題を解決するために、以前のビデオ内に置き換えられる新たなビデオセクションを閉じたピクチャのグループ(GOP: group of pictures)としてラベル付けする方法がある。MPEGエンコードにおいて、閉じたピクチャのグループは、第1のコード化されたIフレーム直後のBタイプフレーム内に用いられる予測が、後方予測のみを使用することを示す。または、MPEGは「broken_link」と呼ばれるフラグの使用を可能とする。このフラグは設定されたとき、第1のコード化されたIフレーム直後の第1のBフレームが、予測に用いられる参照フレームを利用できないために、正しくデコードされないことがあるということを示す。そのデコードは正しくデコードできないフレームの表示を防ぐためにこのフラグを用いる。

【0006】

【発明が解決しようとする課題】 上記解決策は、エンコードされたビデオを劣化させ、品質の低いビデオを提供することがある。

【0007】 そこで本発明は、すでにエンコードされたデータセクションが、エンコードされた他のデータセクションに置き換えられるシステムであって、目にみえるデコード副成分の少ないまたは無い新規な方法および装置を提供することを目的としている。

【0008】 また本発明の目的は、置き換えるべきエンコードされたビデオセクションが、既にエンコードされたビデオ内にエラーすなわちビデオデコード副成分を含めないように、エンコードされた他のビデオセクションと置き換えられるビデオエンコードシステムを提供することである。

【0009】 更に本発明は情報を新規で有用な方法で格納するデータ構造を有するメモリを提供することを目的とする。

【0010】

【課題を解決するための手段】上記及び他の目的はエンコードされたビットストリームのセクションを置き換える新規な方法およびシステムにより達成される。エンコードされたビデオのセクションを取り除くために、デジタル的にエンコードされたビットストリーム内で、エンコードされたビデオの始点および終点を判断する必要がある。これは、目的の画面に対するビットオフセットを判断するために既にエンコードされた各フレームに用いられたビットの数を加算することにより行われる。他の方法としてエンコードされたビットストリーム内の特定フレームの正確な位置又は時間を示す情報を維持するためにディレクトリを用いることができる。

【0011】オリジナルビデオの例えば第2セクションの1/2である1つのGOPの第1セクションは、エンコードされる新たなビデオの始点に位置し、次のGOPの最初の2つのBフレームを正しくエンコード及びデコードするのに必要な切断点の前のGOPの最後のPフレームを生成するために用いられる。この短い初期セクションからのエンコードされたデータは、すでにエンコードされたデータ内で置換されることはなく、スタート点後の最初の2つのBフレームにより参照される参照フレーム（最後のPフレーム）を生成するためにのみ利用される。

【0012】又、前記オリジナルビデオの短いセクションはビデオの終点に位置し、最初にエンコードされたときに用いられた量子化値を用いてエンコードされる。GOP内のエンコードされたビデオのフレームは既にエンコードされたフレームに戻って参照されるので、最初のエンコードのフレームが前記終点直前のビデオの新たなフレームに戻って参照されると、デコードに関する突然の不具合が発生することがある。従って、オリジナルビデオがオリジナルフレームに同等なフレームに戻って参照されるのが望ましい。オリジナルビデオの1/2が編集セグメントの始点で再びエンコードされるが、これは本質的なものではない。

【0013】他の方法として、置換ビデオ内のフレームが置換ビデオの外側のフレームに依存しないように、置換ビデオは周知の閉じたピクチャのグループモードを用いてエンコードされる。この方法の新規な特徴は、予備ビットが双方向フレームに割り当てられ、Bフレームのビット数がPフレームに用いられるビット数の概算値を示すように、この双方向フレームは閉じたGOPモードのために単一方向予測のみを有することである。この方法では、閉じたGOPモードは大きなデコード副成分をBフレームに含まず、BフレームはほぼPフレームと同一の品質を有する。

【0014】デジタル格納媒体に格納されたエンコードされたビデオも本発明の一部を構成する。本発明は又、新規で有用なデータ構造を格納するメモリを含む。このメモリは好適な光ディスクである。

【0015】

【発明の実施の形態】以下、この発明の実施の形態を説明する。

【0016】本発明は、複数の構成要素および機能を備えるデジタルビデオエンコードシステムである。システムのさまざまな構成要素をさらに容易に識別する目的で、本明細書は、以下に示す項に編成されている。

I. はじめに

II. システム構成要素

III. 汎用システム動作

IV. システム代替策

V. 最適エンコード効率を維持しながらオーディオエンコードの数を低減する

VI. エンコードされたビデオの情報を記憶するためのログファイル、ならびにログファイルおよびエンコードされたビデオにアクセスするためのログインタブリタユーティリティ

VII. エンコードされたビデオの品質の変更

A. 期間での品質の変更

B. フレームの領域内での品質の変更

VIII. 品質変更後のビデオの再エンコード

A. 品質が単独フレーム内で変更された場合のビット割当て制御

B. 編集ポイントでのデコードを回避しながら再エンコードする

1. 最初にエンコードしたときに使用された状態にエンコードを復元する

2. ピクチャのグループのクローズモードでの再エンコード

IX. エンコードに使用されるレート量子化モデリング

X. データ構造を使用して、オーディオ、ビデオ、およびサブピクチャのデジタル信号を結合する

I. はじめに

開示されたシステムは、デジタル圧縮システムとしても知られるデジタルオーディオ映像エンコードシステムに関する。本明細書に開示される考えは、あらゆる種類のエンコードスキームで使用可能で、特に、正式にはISO/IEC 13818-2、MPEG-2システム情報を指定するISO/IEC 13818-1、およびMPEG-2の前身であるMPEG-1 (ISO/IEC 1172-ビデオ) として知られるMPEG-2 (エムペグ: Moving Picture Experts Group) に適用できる。本明細書では、MPEG-1およびMPEG-2は集合的に「MPEG」と呼ばれる。MPEG規格は、圧縮されたまたはエンコードされたビデオストリームの構文だけを指定しているが、圧縮をどのように実行するかは指定していない。MPEGエンコードされたデータストリームは圧縮済みのデータストリームであるため、エンコードプロセスは圧縮プロセスと呼ぶことができるし、その逆も可能であることに注意すべきである。また、システムはドルビー (登録商標) AC-3を使用して実行できるオーディオのエンコード、つまりIS

0/IEC 13818-3に定義されるMPEG-2規格にも関する。前記規格のそれぞれが、それらを参照することにより本明細書の一部をなすものとする。

【0017】MPEGビデオ圧縮は、時間的な因果関係の予測圧縮である。MPEGエンコーダは新しいビデオピクチャを受け取ると、新しいビデオピクチャを、エンコーダが既にエンコードされたピクチャに基づいて作成した予測されたピクチャと比較する。予測が以前に伝送された情報に基づいて立てられるために「因果関係(causal)」予測コーディングという用語が使用され、予測が時間的なピクチャ単位で実行されるために、「時間的(temporal)」因果予測コーディングという用語が使用される。予測手順には、ビデオピクチャ内での領域の運動の予測が必要である。したがって、時間的予測は、動き補償予測と呼ばれることもある。

【0018】旧ピクチャと新ピクチャの比較の結果は、後のビデオピクチャに格納される新しい情報を表す。それから、「残留情報(residual information)」と呼ばれる新しい情報は、一連の情報圧縮プロセスを受けることとなる。まず、離散コサイン変換(DCT)と呼ばれる一次数学変換が実行される。DCT演算により、ピクセル残留情報が係数の列に変換される。MPEG圧縮においては、DCTは8x8ピクセルから成るブロックで実行される。同様に、DCT係数も8x8配列数内に入れられる。その後、これらの係数は、量子化ステップサイズつまりq-レベルに基づき、精度の設定された程度まで個別に量子化される(すなわち四捨五入される)。量子化プロセスは、通常、値ゼロが設定された多くの係数を生み出す。さもなければ、この量子化プロセスは、量子化雑音と呼ばれるコーディング構造が生じさせることとなる。

【0019】量子化された係数は、その後、ゼロ値が設定された係数でのランレングスコーディング、およびそれぞれが結果として生じるランレングス値とゼロ以外の係数値をそれぞれ表す2つの数の組の可変長コーディングの組み合わせを使用してコーディングされる。これらのランレングス非ゼロ値の組のコードは、他のオーバーヘッド情報とともに、ビデオピクチャに対応する圧縮済みデジタルデータストリームを形成する。出力データ転送レートや記憶容量のような物理的な制限のために、適切な量の圧縮が達成されるように、量子化の程度を変化させることが必要である。これがバッファフィードバックと呼ばれる。

【0020】MPEGに準拠するデコーダは、エンコードされたデジタルデータストリームを読み取り、エンコードプロセスと逆の演算を実行する。

【0021】一般的には、デジタルデータストリーム中のビット総数は、実質上、ビデオピクチャを表すために必要とされるビット総数を下回るため、圧縮が達成される。しかしながら、MPEGビデオコーディングによって例証されるような圧縮が、決して歪みや人為構造(artifac

ts)を免れないことを認識することが重要である。前述したように、圧縮人為構造の主要なソースは量子化プロセスである。

【0022】開示されたシステムは、圧縮プロセスまたはエンコードプロセスの間に量子化の程度を動的に調整することを目的とする。特に、調整は人間のオペレータによって出される指示に従って行われる。調整は、より大きなビデオシーケンスの部分に適用する場合もあれば、ビデオシーケンス全体に適用することもある。

II. システム構成要素

ここで、類似した参照番号が複数の図を通して同一の部分または対応する部分を示し、さらに特定するとその図1Aを示す図面を参照すると、エンコーダシステムアーキテクチャのブロック図が示されている。ビデオエンコードプロセスとは、中断やエラーなくエンコードおよび編集を達成するために同期して動作する多くの異なった装置を必要とする1つの複雑なプロセスであることに注記すべきである。しかし、オーディオエンコードおよびビデオエンコードは、適切なタイムコードが使用される限り、別個に実行できる。図1Aには4つのワークステーションが描かれているが、通常、エンコードシステムのユーザまたはオペレータが対話するのは1台のキーボードおよびワークステーションだけである。ワークステーションは、プロセッサ、RAM、ROM、ネットワークインタフェース、ハードディスク、キーボード、モニタ、およびそれ以外の従来の構成要素のような従来の汎用コンピュータ構成要素を具備する。

【0023】ワークステーション30は、実施例ではSun SPARC20 Unixワークステーションであるシステム制御ステーションとして機能する。ワークステーション30には、マウスやそれ以外のポインティングデバイスであるキーボード32およびポインティングデバイス33のような従来のワークステーション入力装置だけではなく、グラフィックコンソールディスプレイモニタ31も具備される。ユーザは、おもに、コンソールディスプレイ31に表示されるグラフィックユーザインタフェース(GUI)を介してワークステーション30を視覚的に操作することとなる。

【0024】ワークステーション30は、以下に説明する別のワークステーション40、10および20に接続され、それらを制御する。ワークステーション30は、やはり以下に説明するオーディオインタフェース装置72およびデジタルビデオキャプチャ記憶装置60にも接続される。ワークステーション30は、米国電子工業会(EIA)規格RS-232によって定義される直列インタフェースを介してワークステーション40に接続される。同様に、ワークステーション30とオーディオインタフェース装置72の間の接続も直列RS-232規格による。ワークステーション30とデジタルビデオキャプチャ記憶装置60の間の接続は、業界規格のS-バスインタフェースを介している。

【0025】フレームジョーガー41は、オペレータが、VTR51または52どちらかのビデオを編集を実行しなければならない箇所に位置決めできるようにするために、ワークステーション40に接続される。システムは、ビデオの位置を制御するために、オペレータが入力するタイムコードを使用することができる。ただし、ジョグ制御装置は、ビデオを位置決めするためにジョブノブの使用を可能にすることによって、システムのユーザフレンドリ度を高めている。ワークステーション30は、ビデオディスプレイモニタ61で表示されるビデオが、デジタルキャプチャ記憶装置60または録画VTR62、あるいはその両方からのビデオであるのかどうかを制御する。

【0026】ビデオエンコード装置50は、MPEG-1、MPEG-2、またはそれ以外のビデオ圧縮規格に従って圧縮を実行するデジタルビデオエンコーダである。市販されているMPEG-1エンコーダは、ソニー（登録商標）RTE-3000である。MPEG-2エンコーダは、本明細書に参照により取り入れられる1994年12月27日に出願された日本特許明細書6-326435に説明される教示に従って構築することができる。ワークステーション10は、業界規格S-バスインタフェースを介してビデオエンコード装置50を直接的に制御し、ワークステーション10は、イーサネット（登録商標）ネットワークのような標準ネットワーク装置を介してワークステーション30からコマンドを受け取る。エンコードプロセスの間、ワークステーション10は、ビデオエンコード装置50のビデオエンコード制御プログラムのリアルタイム実行をサポートする。実施例においては、ワークステーション10は、Sun SPARC20 Unixワークステーションである。

【0027】ワークステーション10は中央ワークステーション30からのコマンドに応え、圧縮プロセスを開始する。圧縮プロセスの間、ワークステーション10は、ビデオ信号の特定の領域に適用される量子化の程度を動的に制御する。量子化が変化させられる特定の領域は、空間的領域または時間的領域、あるいはその両方に及び、量子化が厳密にどのように制御されるのかは以下に説明する。

【0028】ビデオエンコード装置50への入力ビデオは、再生VTR 51内のテープからである。再生VTR51は、職業用のデジタルビデオ規格、ITU-R 601（以前はCCIR 601ビデオ規格と呼ばれていた）ビデオテーププレーヤである。この規格は、PALシステムだけではなくNTSCシステムにも適用する。再生VTR 51内のテープには、ビデオカメラまたはフィルムカメラによってキャプチャされたか、あるいはコンピュータによって作成されたさまざまな画像および場面を表す電磁情報が記憶されていることが注記される。ビデオエンコード装置50からのエンコードされた出力は、1つ以上のハードディスク21で記憶するためにワークステーション20に転送される。

【0029】ビデオエンコード装置50によりエンコード

および出力されたビデオデータを表示するためには、MPEG-2のようなデジタルビデオデコーダが必要になる。しかし、本明細書の作成時点では、実現されたシステムは、このようなデコーダはシステムの初期開発段階の間は容易に手に入らなかったため、別個のMPEG-2デコーダを実装していない。したがって、ビデオエンコード装置50は、MPEG-2ビデオストリームをワークステーション20に出力するだけでなく、MPEG-2のエンコードデータに同等な復号化されたビデオ画像がビデオエンコード装置50から録画VTR 52へ出力される。録画VTR 52は、D1インタフェースとして知られる業界規格のデジタルビデオ信号コネクタによりビデオエンコード装置50に接続され、ワークステーション40によって制御される。再生VTR 51および録画VTR 52の両方共、AmpexDCT職業用ビデオテープレコーダを使用して適切に実現される。

【0030】オーディオは、例えばドルビーAC-3フォーマットまたは代わりにISO/IEC 13818-3規格に記述されるようなMPEG-2に準拠するフォーマットにエンコードされ、圧縮される。エンコードシステムのオーディオソースは、4つのデジタルオーディオテープ（DAT）プレーヤ、71a、71b、71cおよび71dである。スタジオ品質のDATプレーヤは、ソニーが市販している。DATプレーヤ71a-71dは、ワークステーション40によって業界規格プロトコルRS-422を介して接続され、制御される。DATプレーヤ71a-71dによって出力されるオーディオ信号は、オーディオエンコーダ70a、70b、70c、および70dにそれぞれ入力される。これらのオーディオエンコーダは市販のドルビーAC-3モデルDP 525エンコーダを使用して実現され、オーディオインタフェース装置72によって制御される。通常、オーディオエンコードは指定されたタイムコードで開始、終了する。オーディオインタフェース装置72は、ワークステーション30からその間のRS-232接続上でエンコード開始タイムコードおよびエンコード終了タイムコードのような命令を受け取る。さらに、オーディオインタフェース装置72は、さまざまなデジタル圧縮済みオーディオデータを、ワークステーション20に伝送される単独デジタルストリームに多重化する。ワークステーション20は、オーディオエンコード装置70により出力される圧縮済みデジタルデータに対応する複数のファイルに、入力デジタルオーディオストリームを多重分離する。ビデオエンコード装置50は、そのビデオ同期信号に基づく同期信号をオーディオインタフェース装置72に提供する。同期信号により、オーディオビットストリームは、ビデオエンコード装置50により出力されるビデオビットストリームに同期できるようになる。

【0031】ワークステーション30は、さまざまなオーディオエンコード装置70a-70dのどれかからオーディオデコーダ73へオーディオデータの複製を提供するようにオーディオインタフェース装置72に指示する機能を備える。オーディオデコーダ73は、圧縮済みのオーディオ

を、ユーザがスピーカ75を介してオーディオを聞くことができるように復号化する。どのオーディオエンコード装置をモニタするかについての選択は、キーボード32またはポインティングデバイス33のどちらかによってコンソールディスプレイ31上でのグラフィックユーザインタフェースを通して人間のオペレータが決定する。オーディオインタフェース装置72の構成は、前述の機能を実行する機能を備えたハードウェアである。このハードウェアは、1つ以上のプログラミングされたマイクロプロセッサまたは開示された機能を達成するために配列される従来の回路要素、あるいはその両方である。

【0032】ワークステーション40はスタジオ装置制御ステーションであり、録画ビデオテープレコーダ52とともに、再生ビデオテープレコーダ51およびデジタルオーディオテーププレーヤ71a-71dの両方を制御するために適切なコマンドを発行する。また、ワークステーション40は、デジタルビデオキャプチャ記憶装置60に、適切な時点でビデオのキャプチャを開始するように命令する。ワークステーション40とさまざまなビデオテープデッキおよびオーディオテープデッキの間の接続は、業界規格のRS-422プロトコルである。このプロトコルを使用すると、各テープデッキはワークステーション40に、標準タイムコードフォーマットでその現在のテープロケーションを知らせることができるようになる。ワークステーション40はこのような情報を使用して、ビデオとオーディオの情報が正しく、再生、記録されるように、さまざまなテープデッキを適切に同期させる。この実施例では、ワークステーション40は、DOSおよびカリフォルニア州、モアパークにあるEditing Technologies Corp. Ensemble Proが市販しているソフトウェアプログラムであるEnsemble Proを実行する標準的なIBM互換パーソナルコンピュータである。ワークステーション30がRS-232通信ポートを介してワークステーション40で実行しているEnsemble Proと通信できるように、マイナーな改良をEnsemble Proプログラムに加えたことを注記する。この改良はEditing Technologies, Corp. を介して市販されているか、あるいは過度の実験を行わないでも当業者により実行できるかのどちらかである。

【0033】デジタルビデオキャプチャ記憶装置60は、エンコードシステム内で、以下に示す2つの機能を行う。第1に、この装置は、ビデオ信号に適用されてきた手動編集の便利な「前と後」比較を行うためにそのキャプチャ記憶機能を使用し、第2に、ビデオディスプレイにグラフィックオーバーレイ機能を提供する。このグラフィックオーバーレイ機能は、例えば、空間手動編集フェーズの間に使用され、人間のオペレータが、量子化およびその結果割り当てられたビットの数が変化する関心のある領域を定義できるようにする。

【0034】本発明では、デジタルビデオキャプチャ記憶装置は、カリフォルニア州、マウンテンビューにある

Viewgraphicsの市販されているDRABMベースのオーディオ製品モデルVS6000として実現されている。この装置に入力されるデジタルビデオは、ITU-R-601 (旧CCIR 601) 画像解像度を備える業界規格D1による。デジタルビデオキャプチャ記憶装置60の出力は、ビデオ信号の赤、緑、および青の構成要素を表す3つの別個のアナログ信号を介してビデオディスプレイに接続される。グラフィックオーバーレイ情報は、デジタルビデオキャプチャ記憶装置60によって出力される最終アナログ信号の作成前に、デジタルビデオキャプチャ記憶装置60により処理されるため、他のビデオデータとともにビデオディスプレイモニタ61では人間のオペレータが見ることができる。

【0035】Unix環境で動作しているワークステーション30は、デジタルビデオキャプチャ記憶装置60を使用するビデオディスプレイモニタ61の上にそのグラフィックウィンドウをマップする。これによって、人間のオペレータは、モニタ61に示されるディスプレイのグラフィックオーバーレイ面で（矩形の描画のような）グラフィック動作を実行できるようになる。装置60は、オーバーレイ機能を実現する。オーバーレイは、メモリ内の情報をモニタを駆動するアナログビデオ信号に変換する前に、ビデオメモリで実行される。システムが使用するVS6000は、ビデオメモリの上部に8-ビットメモリを持つ。この8-ビット面で、コンピュータはあらゆるカラーグラフィックスおよびテキストを描画できる。この8-ビット面での「色」の1つが透明色である。透明色であるこのオーバーレイ面上の任意のピクセルは、その下にあるビデオメモリの値を採る。一般的には、オーバーレイ面の大部分が透明色になり、（線、テキストなどの）グラフィックは透明以外の色を使用する。したがって、ビデオメモリおよびオーバーレイグラフィックメモリの両方からなるディスプレイメモリが形成される。モニタ61で表示するために究極的にはアナログ信号に変換されるディスプレイメモリである。人間のオペレータは、通常、キーボード32と結び付いたポインティングデバイス33 W使用し、このようなグラフィック動作およびこのようにして作成されたグラフィックオブジェクトの操作を実行する。

【0036】ワークステーション20は、オーディオインタフェース装置72だけではなくビデオエンコード装置50からデジタルデータストリームを受け取る。ワークステーション20は、イーサネット接続でワークステーション30に接続され、ワークステーション20は、業界規格のS-バスインタフェースを介してビデオエンコード装置50およびオーディオインタフェース装置72にも接続される。受け取られたデジタルデータストリームは、別個のデータファイルとして1つ以上のハードディスクの中に記憶される。別個のダイレクトメモリアクセス (DMA) カードは、キャプチャされたデジタルストリームデータをハードディスク21に転送するときに使用される。結果的に大量のデジタルデータが生じる長いビデオシーケンスの場

合、ハードディスク21は、Sun Microsystems製のSPARCstorage Array X655A/G5のような市販されているハードディスクを使用して実現できる。この実施例においては、ワークステーション20は、Sun Microsystems製の市販されているSPARKserver 1000である。

【0037】また、ワークステーション20は、ビデオ圧縮済みデータおよびオーディオ圧縮済みデータに対応するデータファイルも、復号化装置への移送または伝送に適切な単独フォーマットファイルにフォーマットする。最終的な媒体はディスクベースであるのが望ましく、その場合フォーマットは、関連するディスクフォーマットに準拠することになる。通常、フォーマットされたファイルは、デジタルテープのようななんらかの中間デジタル記憶手段により、1つまたは複数の最終伝送媒体への転送が発生する機構へ移送される。図示されているデジタル記憶装置22は、このような目的に使用される。本実施例では、デジタル記憶装置22は、市販されている8 mm Exabyteテープドライブを具備する。デジタル記憶装置22のインタフェースは、業界規格の小型コンピュータ用周辺機器インタフェース(SCSI)である。

【0038】複数のSCSI装置を同じSCSIバスに接続できることは既知である。したがって、デジタル記憶装置22は、市販されているデジタルリニアテープ(DLT)ドライブ、磁気光学(MO)ディスクドライブ、またはフロッピー(登録商標)ディスクドライブのような他のSCSI装置も具備できる。これらの構成は容易に対応され、フォーマット段階の前またはフォーマット段階の間に必要となる可能性がある補助データ用の柔軟な入力機構を提供するのに役立つ。例えば、映画の製作では、字幕情報は、大部分の場合、前記のSCSI装置の1つを必要とする媒体内で利用できる。このような場合、補助データはワークステーション20で実行中のソフトウェアプログラムによって読み取られ、補助データがフォーマットされたファイルに格納されるように処理される。補助データは、米国のクローズキャプションフォーマットである字幕データを格納する場合がある。補助データには、プログラムコードやポストスクリプトデータのような希望のデータを入れることができる。

【0039】図1Bは、図1Aのシステムに対応し、図1Aのシステムと同じように機能するオーディオ/映像エンコードシステム98がブロック図で示している。図1Bでは、エンコードシステムへのビデオ入力がカメラ80から供給され、オーディオはマイクロフォン82から供給される。映像情報およびおそらくオーディオ情報を記憶する映画フィルムは、フィルムを照明するためのライトおよびフィルムから画像をキャプチャするための光検出器を具備する転送装置88を介してエンコードシステム98に転送される。入力装置80、82、および88からの情報は、オーディオテープレコーダまたはビデオテープレコーダによってのような、エンコードシステム98によってエンコード

される前に電磁フォーマットで記憶される。オーディオテープおよびビデオテープは、図1Aのシステムによりエンコードされる。

【0040】エンコードシステムにより作り出される最終的なデジタルビットストリームは、光ディスクをプレス加工する光ディスクマスタリング装置90、光ディスクに書き込む光ディスク書込み装置84に送られるか、あるいはテレビジョン送信機86を介して、テレビジョン受像器またはセットトップボックスデコーダに伝送されることがある。光ディスクのプレス加工および作成は、例えばマスタディスクを使用する既知で市販されている方法で実行される。エンコードされた情報は、ビデオオンデマンドシステムでも使用できる。エンコードされたビットストリームは、究極的には、エンコードプロセスに対応する復号化プロセスにより復号化され、オーディオ映像情報はテレビまたはモニターでユーザに表示される。復号化装置は、テレビまたはモニターで表示するためにエンコードされたディスクを復号化するデジタルビデオディスクプレーヤを具備する。

【0041】エンコードシステムがオーディオおよび画像を表す電気信号を入力し、電気信号を新しいフォーマットに変換し、究極的には、エンコードされた元の情報を作成し直すために信号が復号化されることは、図1Bから明らかである。

III. システム動作

図2は、図1にブロック図で示されるシステムの一般的な動作を説明するフローチャートである。図2から分かるように、システムはその動作サイクルの間に複数の「状態」を経過する。濃い実線は標準的な操作経路を示し、実線は中断終了経路を描き、破線はバイパス経路を示す。

【0042】システムによる最初の動作はセットアップ状態100で発生する。このステップによって、圧縮状態102が入力される前に、入力パラメータを指定することができる。圧縮状態102では、再生VTR 51内のテープに記憶されるビデオ画像が、ビデオエンコード装置50によりデジタル圧縮されたフォーマットにエンコードされる。

【0043】圧縮状態102の後、ユーザは、編集状態104の過去に圧縮されたデータの画質に基づいて手動で圧縮プロセスを編集することができる。これによって、ユーザは、ビデオの1フレームの領域内または期間で画質を改善または低下させることができる。「編集」という用語が、場面自体をピクチャストリームに付加したり、削除することを意味しないことに注意すべきである。本明細書中では、「編集」という用語は、量子化および続いてビデオの質を変更することを意味する。画像が編集状態104で編集された後、新しいユーザ指定パラメータに従ってデータをエンコードするために、圧縮状態102で再び圧縮動作を実行することが必要である。

【0044】ユーザがビデオの編集された形式を確認し

てから、字幕のようなオーディオ、ビデオおよびそれ以外の情報を含むすべての情報が、フォーマット状態106で、まとめて希望のフォーマットに結合される。完了状態108では、あらゆる一時ファイルが消去され、プロセスは終了状態110で終了する。図2に説明される個々の状態に関するさらに詳細な情報は、ここでは図3から図6を参照して説明する。

【0045】図3は、セットアップ状態100で実行されるステップを示している。ステップ120は、セットアップ状態を開始し、ステップ122はシステムを初期化した。このステップでは、ワークステーション30に格納されるシステム制御ソフトウェアの実行が開始し、コンソールディスプレイ31に表示され、キーボード32およびポインティングデバイス33を介してパラメータを入力させるユーザインタフェースを担当するグラフィックユーザインタフェース(GUI)プログラムのようなそれ以外のプログラムが起動される。また、ステップ122の間に、ワークステーション30のシステム制御装置ソフトウェアが他の装置に照会し、システムの利用可能な装置を含むシステム状態を突き止める。

【0046】ステップ124では、GUIを通してユーザに情報が表示される。歓迎のメッセージ、システムに関する情報、エンコードプロセスに使用されるパラメータ、および過去にエンコードされたあらゆるファイルの存在が表示される。表示されたシステム情報には、オーディオエンコーダ、VTR、オーディオテープデッキ、およびその構成を含むシステムに接続される装置を具備する。システムに既にエンコードされたファイルが格納されている場合は、情報はファイルを作成するのに使用されたオーディオエンコードパラメータおよびビデオエンコードパラメータを記述する情報がユーザに表示される。

【0047】ステップ126では、それから、標準システムパラメータまたはユーザ指定システムパラメータを使用するエンコード、システムパラメータに対する変更、または選択されたパラメータが受入れ可能であることの確認を開始するためのコマンドのようなユーザからの情報を入力する。

【0048】それから、ステップ128では、エンコード動作と復号化動作の準備をするために、グラフィックユーザインタフェースからシステム制御装置ソフトウェアに、入力されたデータを転送することにより、入力データを処理する。セットアップ状態は130で終了する。

【0049】図4には、圧縮状態102で実行されるプロセスがブロック図が示される。標準動作フローに従い、ユーザは、既に、セットアップ状態で圧縮されるビデオおよび付随オーディオシーケンスを指定した。それから、ユーザは、システムに、自動手順、つまり自動圧縮手順を実行するように指示する。自動圧縮手順の間のビデオ圧縮により、ユーザは期間で、あるいはビデオの1フレームの中で動的量子化プロセスを制御することはできな

い。これは、コーディング決定が、ビデオエンコード装置50により計算される客観的な基準に基づいて下される初期圧縮実行として意図される。自動圧縮がすでに実行されていた場合は、ユーザは、空間的（フレームまたはビデオ上で）または時間的（ビデオシーケンスで）のどちらかで、編集圧縮モードでの量子化を手動で調整できる。

【0050】圧縮状態では、ステップ140で開始した後、ステップ142で圧縮動作に備える。この動作中、ビデオのタイムコードおよび圧縮パラメータとともに圧縮される付随オーディオのような圧縮プロセスのパラメータが、ワークステーション10を介してワークステーション30からビデオエンコード装置50にダウンロードされる。

【0051】それから、ステップ144では、希望の圧縮の種類が決定される。初めて情報を圧縮するときには、圧縮プロセスは自動的に、ユーザの介入なく実行される。ステップ146では、自動圧縮プロセスの準備をする。この準備中に、ワークステーション30は、デジタルオーディオテーププレーヤ71a-71dおよび関連するオーディオエンコード装置70a-70dの内のどれを使用しなければならないかを決定する。また、ワークステーション30は、ワークステーション20に、オーディオインターフェース装置72を介した、ビデオエンコード装置50およびオーディオエンコーダ70a-70dの出力からのビットストリームキャプチャに備えるように命令する。さらに、再生VTR 51、録画VTR 52、およびDATプレーヤ71a-71dは、テープを適切な開始位置まで進める。それから、ワークステーション40が再生ビデオテープレコーダ51、録画テープレコーダ52、およびDATプレーヤ71a-71dに、ステップ148での移動を開始するように信号を送る。その後、ステップ150で、ビデオエンコード装置50を使用するMPEG-1またはMPEG-2フォーマットのような希望のフォーマットに従ってビデオデータがエンコードされる。また、オーディオデータも、ドルビーAC-3フォーマットのような希望のフォーマットに従ってエンコードされる。手動編集プロセスを後で補助するためにログファイルを作成する必要がある。ログファイルは、ビデオデータの後の編集を実行するために必要となるエンコードされたビデオデータのパラメータを示す。エンコードプロセスは、適切なタイムコードに到達するステップ160で停止する。このタイムコードは、エンコードされるべき情報の最後である場合もあれば、ビデオシーケンスの最後ではないユーザが指定したポイントである場合もある。

【0052】ステップ144で、自動圧縮プロセスがすでに実行され、ユーザが、以下に図5に関して説明するように、編集状態の間に指定されたパラメータを使用してビデオの1つ以上のフレームの画質を手動で変更することを希望すると判断されると、ステップ144で、編集圧縮動作を実行する必要があるかどうか判断されるだ

ろう。ステップ152では、手動編集が開始するポイントである希望の開始ポイントへの再生VTR 51の合図を含む編集圧縮動作の準備が行われる。また、オーディオ情報は、いったん自動圧縮動作でエンコードされると、変更する必要はないので、オーディオ装置はディスエーブルされる。それから、ステップ154で、再生VTR 51および録画VTR 52が移動する。

【0053】ステップ156では、その後、図5に説明する編集状態でユーザによって指定されたパラメータに従ったビデオデータのエンコードが開始する。編集圧縮は、終了タイムコードに達したステップ160で終了する。ステップ162では、ユーザに、圧縮およびエンコードが完了した旨のメッセージが表示され、プロセスはステップ164で終了する。

【0054】図5は、編集状態104で実行されるプロセスを説明する。前述したように、ここで言われる編集とは、ビデオシーケンス内で場面を削除、短縮、または移動させる従来のビデオの編集ではない。本明細書でいう編集状態とは、ビデオの一部またはビデオの時間シーケンスの質を、手動でビデオの量子化を設定することにより変更することである。

【0055】ステップ170で編集状態を開始した後、ステップ172では、ユーザはエンコードされたビデオのビット割当てを変更することによるビデオの手動の編集を希望するかどうか尋ねられる。ユーザがビット割当ての変更を望まない場合、プロセスはステップ188で終了する。ユーザがビット割当ての変更を望む場合は、ユーザはステップ174で編集対象のビデオセグメントを定義する。これは、画質を変更しなければならないビデオの期間を選択することによって実行される。それから、ステップ176で、ユーザが空間的編集または時間的編集を希望するのかを判断する。空間的編集がビデオのフレームの中でビットの量子化または割当てを変更するのに対し、時間的編集は、ビットの割当てを期間で変更する。ユーザが空間的編集を選択すると、フローは、編集されるフレーム内の領域を入力するステップ178に進む。また、適用される変更の相対的な程度も入力される。本発明の場合、ゼロを含む-5から+5の整数目盛りが、変更の相対量を示すために使用される。ユーザは、キーボード32またはポインティングデバイス33あるいはその両方を使用してビデオの領域をマークし、-5から+5の間で-5と+5を含む整数の内の1つを指定する。ユーザは、ある領域に以前設定された量子化レベルを変更してはならないことを示すこともできる。例えば、ユーザはある一定の領域のビット割当ての増加を希望する場合、それ以外の領域のビット割当てを減少しなければならない必要がある。ユーザがある領域を「保護つき(protected)」とも呼ばれる変更不可と設定すると、画質領域の改善が必要となるビットは保護つきの領域からは採取されない。

【0056】ステップ176でユーザが時間的編集を希望

すると判断されると、フローは、どのように期間を編集しなければならないのかを入力するステップ180に進む。空間的編集と同様に、時間的編集の場合も、ユーザは、ビデオのある特定のシーケンスに適用される変更の相対度を示す-5と+5の間の整数値を指定する。この変更は、選択された期間全体で有効となる。

【0057】空間的編集または時間的編集のどちらかがステップ178と180で実行されてから、ステップ182で、ビデオシーンが、手動で入力されたパラメータに従って、エンコードし直され、例えばMPEG-2フォーマットに圧縮し直される。ユーザが画質がどのように変更されたのかを詳しく比較できるように、既にエンコードされたデータは新規に編集されたデータの隣に表示される。

【0058】オペレータが編集対象のセグメントを示すと、VTR 52からのビデオが記憶装置60に転送され、記憶装置60で利用可能な総メモリの多くても半分を占める。そのセグメントは、「前の」セグメントに対応する。編集-圧縮プロセスの間、ビットストリームはワークステーション20によってキャプチャされ、マスタビデオビットストリームファイルから別個のビットストリームファイルとして記憶される。マスタビデオビットストリームファイルは、映画全体の圧縮済みビットストリームである。編集-圧縮されたビデオはVTR 52で録画される。エンコードプロセスの間、オペレータはピクチャを、それがVTR 52で録画されているのと同時に表示することになる。オペレータがエンコードした後のピクチャの表示を希望する場合は、VTRを再生機械として使用できる。オペレータが「前と後」の比較を実行したいと考える場合、VTR 52の対応するセグメントが記憶装置60に転送され、装置60の総メモリの残りの半分までを消費する。

【0059】このエンコードの後、オペレータは、変更されたビデオを保存する必要があるかどうかを決定する。オペレータが「後」のセグメントを無視することを決めた場合、マスタビットストリームは変更されずに残され、編集-圧縮済みのビットストリームは削除される。VTR 52がマスタビットストリームを反映できるようにするには、「前」のセグメントを録画する目的だけに、編集済みのセグメントで別の自動-圧縮が実行される。装置60を「前」のセグメントをVTR 52に転送して戻すために装置を使用するという代替策も可能である。オペレータが「後」のセグメントを記憶することを決めた場合、以下に説明するように、マスタビットストリームファイルを更新して、編集-圧縮済みのビットストリームファイルを入れる必要がある。

【0060】ユーザが、ステップ184でビデオが受け入れられると判断した場合、既にエンコードされたデータは新規にエンコードされたデータで置き換えられる。このステップは、以下の別項でさらに詳しく説明する。新規にエンコードされたデータが以前エンコードされたデータを置き換えると、ステップ192で、ユーザが更に編

集を希望するかどうかを判断する。さらに編集を希望する場合には、フローはステップ174に戻る。ユーザがビデオの編集を終了したら、フローは編集状態を終了するステップ188に進む。

【0061】ユーザが、ステップ184でビデオが受け入れられないと判断した場合は、ユーザは、編集変更が記録されないようにステップ186で編集動作を終了するか、あるいはユーザはビデオを編集し直す。ビデオは、編集で新しいパラメータを選択するユーザによって編集し直される。ユーザが希望する場合は、編集対象の新しいビデオセグメントを定義することも可能である。

【0062】ユーザが圧縮（自動圧縮だけ、または編集圧縮とともに自動圧縮のどちらか）に満足すると、圧縮済みビデオ、圧縮済みオーディオ、およびそれ以外のすべてのファイルを含む個々のデータファイルが、最終フォーマットへのフォーマットの準備が完了した状態となり、フォーマット状態166に入る。補助データには、最終的なフォーマット済みデータファイルに記憶される字幕のような情報を含むことがある。図6にブロック図で示されるフローチャートでは、ステップ200でのフォーマット状態の開始後、ユーザが補助データを記憶することを決定すると、ワークステーション20は、フロッピーディスクドライブ内にフロッピーディスクなどを具備するデジタル記憶装置22から補助データを読み取る。それから、ステップ204で補助データがエンコードされたオーディオおよびビデオと結合され、最終的なフォーマット済みファイルを作り出す。システムによって作成される最終的なデジタルビットストリームは、ISO/IEC 13818-1に定義される「プログラムストリーム」として構築できる。代わりに、最終的なビットストリームは、直接衛星DSSフォーマット、光ディスクでの使用に適したVBR（可変ビットレート）またはその他の種類のフォーマットのような任意の種類のフォーマットとすることができる。エンコードされたビデオとオーディオは別個に記憶できるため、システムは、同じエンコードされたビデオとオーディオを別々の最終フォーマットにエンコードすることができる。これは、さまざまなフォーマットを作り出す機能を備える1つのフォーマットを使用するか、または別個のフォーマットによって達成される。ステップ206では、フォーマットされたファイルがディスク21で記憶される。

【0063】フォーマットされたファイルがステップ210のデジタル記憶装置22内のテープに書き込まれるステップ208で完了状態になる。その後で、もはや必要とされていない一時ファイルを削除するさまざまな「ハウスクリーニング」機能を実行するステップ122で終了状態に入る。圧縮プロセスが終了される前のような異常終了要求が発生した場合、打ち切り手順が、オーディオテーププレーヤおよびビデオテーププレーヤを停止し、破壊されている可能性があるデータファイルもクリーンアップ

するワークステーション30で実行される。プロセスは、それからステップ216で終了する。

IV. システム代替策

望ましいシステム構成要素およびその動作は前記に説明したが、言うまでもなく代替ハードウェアが図1Aおよび図1Bに開示されるハードウェアを代用することができる。ビデオエンコード装置50は、利用できる場合にはMP EG-2ビデオエンコーダであるのが望ましい。しかし、MP EG-1ビデオエンコーダとしても知られる市販されているISO/IEC 1172-ビデオ標準エンコーダを使用することもできる。MPEG-1エンコーダに対する必要な改良は、MPEG-1バッファ管理モジュールが、ワークステーション10のような外部ソースから制御できるようにする必要がある。さらに、入力ビデオサポートは、ITU-R 601（前CIR 601）解像度ビデオを格納するために増大されなければならない。このような改良は、過度の実験を行わないでもデジタルビデオエンコードの当業者により実行できる。

【0064】4つのオーディオテーププレーヤと4つのオーディオエンコーダが図1Aに描かれているが、それ以外の構成はオーディオシステムに容易に採用できる。例えば、1台の市販されている8 mmデジタルオーディオプレーヤを複数のDATプレーヤの代わりに使用することができる。1台のTASCAMデジタルオーディオプレーヤを使用すると、ワークステーション40から必要とされるのは1つのRS-422制御信号だけであるが、最大8つの別個のオーディオチャンネルを同時にサポートできる。オーディオエンコーダが同時に8つのオーディオ入力を受け入れることが可能でなければならない、オーディオインタフェース装置72もオーディオデータストリームの4つから8つへの増加に対処するために改良されなければならないのは言うまでもない。

【0065】オーディオおよびビデオの同時エンコードの代替策として、オーディオおよびビデオエンコードは、別個の時間またはロケーション、あるいはその両方で実行され、後で最終的なフォーマットに結合することができる。これには、SMPTE（全米映画テレビジョン技術者協会）タイムコードフォーマットのような既知のタイムコードの使用が必要となる。タイムコードは、ビデオテーププレーヤ、ワークステーション、または別個のスタンドアロン式タイムコード作成プログラムによって作成される。

【0066】録画VTR 52は、ビデオデコーダが利用可能で、ワークステーション20に接続されている場合には排除できる。その場合、再構築されるビデオは、圧縮フェーズ中のビデオエンコード装置50から録画される代わりに、ディスク21内のディスクファイルから作成されるだろう。録画VTR 52を排除すると、テープ費用だけでなく装置費用という点でも、システムの費用が大きく節約される。

【0067】さらに、グラフィック動作をディスプレイビデオのグラフィックオーバーレイ面で実行する能力は、市販のX-端末オプションを使用して、コンソールディスプレイ上でサポートすることができる。例えば、コンソールディスプレイ31は、ビデオをオーバーレイグラフィックウィンドウで表示する機能を備えたX-端末で置き換えられる。ペンシルバニア州のKing of PrussiaのHuman Designed Systems社のHDSV View Stationのような市販されているマルチメディアディスプレイ端末は、表示および手動編集領域定義動作のためにX-端末に接続することができる。ただし、ビデオデコーダの職業用品質のビデオは、人間のオペレータが再構築されたビデオ信号の質を評価できるように、図1Aに図示されるモニタ61のような職業用モニタで表示しなければならない。

V. 最適エンコード効率を維持しながらオーディオエンコードの数を低減する

前述したように、図1Aは、マルチパスビデオエンコード装置50および4台のワンパスオーディオエンコード70a、70b、70c、および70dを備えたエンコードシステムをブロック図している。再生VTR 51は、ビデオエンコード装置50にエンコードされるべきビデオ信号を供給し、デジタルオーディオテープ(DAT)プレーヤ71a-71dが、それぞれワンパスオーディオエンコード70a-70dにエンコードされるべきオーディオ信号を供給する。

【0068】本発明の実施例では、最終フォーマット済みオーディオ映像データストリーム内の8つのオーディオトラックをエンコードする。各オーディオトラックは、1つ以上のオーディオチャネルを具備する。オーディオトラックは他の装置でもオーディオを格納できるが、例えば、あるオーディオトラックが左チャネルと右チャネルを備える場合がある。図7に説明する。もっとも効率的なオーディオ映像エンコードを実行するためにオーディオエンコードの数を求める方法、およびこのオーディオ映像エンコードを実行する方法は、図7に説明する。図7では、開始後、ステップ240でビデオエンコードに必要とされるパスの数Pが求められる。ビデオをMPEGフォーマットにエンコードするためには、通常、2つ以上のパスがビデオのエンコードに必要となる。本発明の場合、望ましいパス数は2であるが、3つ以上のパスも可能である。それから、ステップ242でエンコードされるオーディオトラックの数Tが求められる。本発明では、異なったオーディオトラック数も可能であるが、オーディオトラックの数は8が望ましい。次に、ステップ244では、もっとも効率的なオーディオ映像エンコードを実行するために必要とされるワンパスオーディオエンコード数AEが計算される。必要なオーディオエンコードの数は、ビデオエンコードに必要とされるパスの数で除算されるエンコードされるべきオーディオトラック数に等しい。実施例では、エンコードの2つのパスで除算される8つのオーディオトラックは、必要な4つのオーディオエ

ンコードに等しい。オーディオエンコードの計算された数が整数ではない場合、オーディオエンコードの計算された数は、整数となるように切り上げる必要がある。例えば、7つのオーディオトラックだけしかツーパスビデオエンコードにエンコードされない場合、 $7/2=3.5$ となり、3.5は次にもっとも大きい整数4に切り上げられる。本システムによって実行されるエンコードプロセスの間、1つのビデオパスはオーディオエンコードの内3台しか使用しないが、それ以外のビデオエンコードパスは4台のオーディオエンコードすべてを使用する。ステップ240-244が、システムの初期設計時にだけ実行される必要があり、オーディオ映像作業がエンコードされるたびに必要とされるワンパスオーディオエンコードの数を計算する必要はないことに注意すべきである。さらに、ステップ240、242および244は、ワークステーション30のようなコンピュータによってコンピュータ/マシン実装されるか、あるいはユーザまたはオーディオ映像エンコードシステムの設計者により判断される。

【0069】オーディオエンコードの数が求められ、オーディオ映像エンコードシステムがセットアップされてから、オーディオ映像エンコードプロセスが開始できる。ステップ246では、1にカウンターがセットされる。カウンターは、ビデオエンコードシステムの各パスをカウントするに用いられる為に可変である。それから、ステップ248は、オーディオトラックのすべてではないが、いくつかでワンパスオーディオエンコードを実行しながら、第1ビデオエンコードパスを実行する。例えば、8つのオーディオトラックおよび1つのツーパスビデオエンコードが存在する場合、第1パス中に4つのオーディオエンコードを使用できる。それからステップ250で、カウンターが1だけ増される。ステップ252では、ステップ248で使用されたのと同じオーディオエンコードを使用して、エンコードされていないオーディオトラックのワンパスオーディオエンコードを実行しながら、次のビデオエンコードパスを実行する。例えば、第2ビデオエンコードパスの間、ステップ249で使用する4つのオーディオエンコードは、オーディオトラック5から8をエンコードできる。それからステップ254で、カウンタがP、必須ビデオエンコードパス数に等しいかどうか判断される。実施例では、2つのビデオエンコードパスしかないため、プロセスは終了するだろう。終了しない場合、フローはステップ250、252、および254から成るループを実行するために戻る。

【0070】言うまでもなく、8つのオーディオトラックがあるオーディオ映像作品用のツーパスビデオエンコードを使用するだけでなく、それ以外の実施例も可能である。例えば、スリーパスビデオエンコードは、6つのオーディオトラックで使用できる。この状況では、オーディオエンコードのもっとも効率的な数は3 ($6/2=3$) である。ツーパスビデオエンコードが使用され、合計6

つのオーディオトラックがエンコードされる場合、オーディオエンコーダのもっとも効率的な数は3となるだろう。代替策として、スリーパスビデオエンコーダを、エンコードするオーディオトラックが5つあるシステムで使用することができる。この代替システムでは、オーディオエンコーダの内の1つしか使用されず、他の2つのビデオエンコードパスが、両方オーディオエンコーダの両方が使用されている間に発生する一方、ビデオエンコードパスの内のどれか1つが動作している状態の2つのワンパスオーディオエンコーダが必要である。従来の技術の当業者に明らかになるように、ビデオエンコーダとオーディオトラック数のこれ以外の組み合わせも、本明細書に記載される教示を考慮すると可能である。

【0071】できるかぎり効率的であるという点での本発明の重要な一面とは、ビデオエンコードのパスが発生している間につねにオーディオエンコードを実行し、オーディオエンコードが実行している間につねにビデオエンコードのパスを実行するということである。この取り決めでは、ビデオエンコード動作の1パスの間に、オーディオエンコーダの少なくともいくつかがつねに使用されることになる。できるかぎり効率的であるためには、オーディオトラック数をオーディオエンコーダの数で均等に除算できない場合（つまり、エンコードされるオーディオトラックの数をビデオパス数で除算すると余りが出る場合）、エンコードプロセス全体で、ビデオパスから余りを差し引いた数に等しい数の休止していたオーディオエンコーダがある。例えば、スリーパスビデオエンコーダおよび4つのオーディオトラックの場合、2つのオーディオエンコーダが必要とされ、余り1が出る（例えば、 $4/3=1$ で余りが1）。したがって、もっとも効率的であり、スリーパスビデオエンコードが完了されのと同時にオーディオエンコードを完了するためには、オーディオエンコーダが休止となるのはわずか2回である（例えば、3つのビデオパスから余り1を差し引くと2に等しい）。エンコーダが休止する2回は、同じパスで発生可能であり（同じビデオエンコードパスに異なった2つのオーディオエンコーダがある）、前記例のそれ以外の2回のビデオエンコードパスのそれぞれがオーディオエンコーダのすべてを活用するか、あるいは1つの休止オーディオエンコーダがある2つのビデオエンコードパスがあるという意味である。これらの2つのパスでは、同じまたは別のオーディオエンコーダが休止となることがある。

VI. エンコードされたビデオの情報を記憶するためのログファイル、ならびにログファイルおよびエンコードされたビデオにアクセスするためのログインタブリタユーティリティ

開示されたエンコードシステムは、当初、MPEG-2、MPEG-1、またはそれ以外のデジタルビデオエンコードフォーマットのような圧縮済みデジタルフォーマットにビデオを

エンコードする。このエンコードは、システム構成要素に関する第II項および汎用システム動作に関する第III項に説明する。したがって、エンコードされたビデオデータは、例えばハードディスク上のデジタルフォーマット内で記憶される。

【0072】デジタルビデオエンコードプロセスの間、ビデオの統計およびそれ以外の情報が作成され、ログファイルに記憶される。MPEGビデオエンコードに精通した従来の技術の当業者は、ビデオを記述する統計および情報の種類、およびこれらの統計および情報を作成する方法を熟知している。従来のエンコードシステムでは、この種類の情報およびデータは作成されるが、通常、ビデオエンコードの後に廃棄される。しかし、本発明の場合、ピクチャをコーディングするために使用された合計ビット、ピクチャをエンコードする平均二乗エラー、ピクチャあたりの平均量子化スケールなどのような作成されたこの情報は、エンコードプロセスの後で、システムRAMからログファイルにダンプされる。本発明によって使用されるログファイルの例は、図8A-10Bに説明する。これらのログファイルは、VBRシステムで特に有効であり、ビデオエンコーダの初期パス（複数の場合がある）の間に作成される。それから、ログファイルは、最終エンコードパスの間およびビデオの後での編集および再エンコードの間に使用される。

【0073】本発明により使用されることがある、シーケンスレーヤのログファイル、ピクチャレイヤログファイル、およびマクロブロックレーヤのログファイルという異なった3種類のログファイルがある。これらのログファイルのそれぞれが、その各レイヤの統計および情報を記述する。図10(A)にブロック図で示される完全フォーマット、および図10(B)にブロック図で示される短縮フォーマットという異なった2種類のマクロブロックレーヤのログファイルフォーマットがある。マクロブロックレーヤの詳細な情報が必要とされる場合には、完全フォーマットが使用され、マクロブロックの詳細のすべてを追跡調査する必要がある場合には、短縮フォーマットが使用される。2時間の映画中のマクロブロックの数は多数であるため、マクロブロックログファイルは大量の記憶スペースを消費する。したがって、大きな記憶スペースが使用できない限り、完全または短縮フォーマットのマクロブロックファイルを記憶することはできない可能性がある。ただし、マクロブロック情報を入手する必要がある場合は、エンコードビデオを復号化するか、あるいはそれ以外の手段でこの情報を再構築することができる。

【0074】図8A-10Bにブロック図で示されるログファイルフォーマットの詳細の多くは重要ではなく、これらのファイルフォーマットは、単に、エンコードプロセスの間に作成できるデータの例図としてのみ提供されたことに注記すべきである。しかし、ログファイル中の情報

のいくつかは、ビデオの質を変更するプロセスの間に使用されるので重要である。以下に説明するように、エンコードされたビデオの編集ポイントを決定するためには、初期のピクチャにより消費されるビット数を総計する必要がある。合計される情報は、図9Bおよび図9Cに描かれ、ピクチャごとのビット数 (generated_bit_picture)、シーケンスヘッダ (bits_sequence_header)、ピクチャのグループ(GOP)ヘッダのビット数(bits_GOP_header)、およびスタッフィングとして使用されるビット数またはスペースを埋めるために使用されるビット数(stuffing_size)に関する。

【0075】図11は、ログファイルを使用して実行されるさまざまな機能のフローチャートを説明する。ビデオセグメントの初期エンコードの後、ステップ270では、図8A-9Cにブロック図で示されるビデオエンコーダの自動実行から作成されたピクチャおよびシーケンスレーヤのログファイルを読み込む。それから、ステップ272でエンコードされたビデオに録画されるピクチャごとのログファイルレコードのバイトオフセットを示すインデックステーブルを作成する。2つのテーブルが作成される。一方はフレーム番号により、もう一方はタイムコードによりインデックスが付けられる。タイムコードインデックステーブルには、ピクチャごとのログファイルの中にオフセットされるバイトが記憶され、タイムコードの昇順で並べ替えられる。同様に、ピクチャフレームインデックステーブルには、(エンコーダによって確認されるように) 入力されたピクチャ番号の昇順で並べ替えられるログファイル内へのバイトオフセットが記憶される。

【0076】作成されたインデックステーブルは、任意の希望されるピクチャまたはマクロブロックのログ情報をすばやく選び出すのに有効である。ログファイル内のシーケンス、ピクチャまたはマクロブロックのデータを記憶するためには固定サイズレコードが使用されているが、エンコーダはそれらをコーディングする前に入信ピクチャを並べ替える。加えて、3:2の割合で毎秒30フレームに引き下げられた映画(毎秒24フレーム)ソースを取り扱う場合、ある種のタイムコードは省略される。インデックステーブルは、ピクチャの並べ替えおよびタイムコードの省略にも関わらず、適切なデータの位置をすばやく突き止めるための手段となる。

【0077】ビデオのさまざまなフレームと期間がその品質を手動で改変されるに従い、エンコードされたビデオは頻繁に変化するもので、本発明によるエンコードシステムが、各ピクチャをコーディングする際に使用されるエンコードされたビット数のディレクトリまたはインデックスを維持しないことを注記すべきである。ただし、エンコードされた情報を記述するログファイルは固定サイズレコードであるため、ピクチャまたは期間の品質が改変されても、ログファイル内の情報のディレク

トリまたはインデックスを維持することは容易である。

【0078】インデックステーブルの作成後、ステップ274を使用することにより、ユーザは、ブロック276、278、および280に描かれるさまざまなログファイルユーティリティを使用するログファイルからさまざまな情報をすばやく入手できる。ログファイル中のレコードが必要とされる場合、作成されたインデックスは、ログファイル内の希望のレコードを突き止めるために使用され、完全なレコードがログファイルから引き出される。それから、このレコードは希望される個々のパラメータを検索するために解析される。個々のレコードをインデックステーブルを使用して解析するプロセスは、他のログ解釈プログラムユーティリティのすべての基礎となる。解析プロセスは既知であり、従来の技術の当業者は、編成されたファイルから希望の情報を検索する方法を理解している。

【0079】ステップ276は、ログファイルのカットアンドペーストユーティリティである。フレーム内でまたは期間で手動で量子化(および画質)を変更した後、ログファイルのオリジナルセットは、編集中に発生した変更に対応するように更新されなければならない。ログファイルのカットアンドペーストユーティリティは、編集済みのセグメントタイムコードに対応するログファイル内でオフセットの位置を突き止め、ログファイル内の情報を再エンコードされたビデオに対応する新しい情報で置き換える。再エンコードのプロセスについては、以下のVIII項に説明する。

【0080】ステップ278では、デジタルでエンコードされたビットストリームの編集を可能とするためにパラメータを入手するためのユーティリティを示している。ビデオが最初にエンコードされ、ユーザがビデオのフレームの品質または期間の改変を希望する場合、削除され、希望の品質を備えた新規にエンコードされたビデオで置き換えられなければならないエンコードされたビデオ内の適切な部分を突き止める必要がある。

【0081】エンコードされたビデオのビットストリーム内の編集ポイントのバイトオフセットは、編集ポイントに対応するピクチャまで各ピクチャをエンコードするのに使用されるビットの累積項目を計算することにより求められる。指定のピクチャが位置するファイルの最初からのバイトまたはビットの総数を求めるために、以前のピクチャのそれぞれをエンコードするのに使用されたビット数が合計される。合計されたピクチャログファイルから入手される情報とは、 i 番目のピクチャをエンコードするのに使用されるビット B_i の数、シーケンスヘッダを構成するビット S_i の数、ピクチャのグループ(GOP)ヘッダをエンコードするのに使用されるビット G_i の数、スタッフィングとして使用されるビット T_i の数である。 N 番目のピクチャのビットオフセットは、次に示すように計算される。

【0082】

$$\sum_{i=1}^{N-1} (B_i + S_i + G_i + T_i)$$

(1)

バイトオフセットは、ビットオフセットを8で除算することによってだけ計算される。エンコードされたビデオビットストリーム中のアドレスオフセットを求めるためにログファイルを使用する代替策として、ビットストリームデコーダを使用して、ピクチャオフセットのディレクトリを作成することができる。ただし、エンコーダは、ディレクトリを更新するために編集が終わるたびにビットストリーム全体を実行しなければならないため、このアプローチは面倒である。

【0083】最後に、ステップ280では、タイムコードをもっとも近いピクチャのグループ (GOP) の最初または最後に四捨五入するためのユーティリティを示している。エンコードされたビットストリームのすべての切取りまたは編集は、ピクチャのグループ (GOP) の境界で実行される。ただし、GOP境界はビデオに編集を実行するオペレータにはわからない。オペレータがいったん編集セグメントの開始ポイントおよび終了ポイントを指定すると、システムはGOP境界を計算する。GOP境界は、各ピクチャレコード内に記憶されるGOP番号と呼ばれるピクチャレイヤパラメータを検討することにより設定される。ピクチャレイヤログファイルの連続レコードを通してトレースすることにより、図9Bの中のGOP番号 (GOP_number) が変化すると、GOP境界があると判断される。

【0084】ログファイルおよびログファイルユーティリティは、ビデオエンコーダに密接に関係しているので、ログファイルは図1Aに図示されるワークステーション10のハードディスク内に記憶され、ログファイルユーティリティはワークステーション10により実行される。代わりに、ワークステーション30またはそれ以外のの任意の処理システムにログファイルが格納され、ログファイルユーティリティが実行されることがある。

VII. エンコードされたビデオの品質の変更

A. 期間での品質の変更

エンコードシステムにより作成されるエンコードされたビデオは、光ディスクのような最終的なデジタル記憶媒体に書き込まれることを目的とする。ディスク上のエンコードされたビデオは、一定ビットレート (CBR) 動作と比較すると、ビデオの異なったセクションの間でのビットレートの大きな変更の可能性を提供する可変ビットレート (VBR) でエンコードされる。VBRコーディングにより、著しく異なるビット数を異なったピクチャに割り当てることができるようになるため、画質は経時的に変化する。さらにVBRは、CBR動作では不可能なスケール内でのユーザによって定義されたビデオのセクションにデータレートを割り当てし直す可能性も提供する。CBRにより

同じことを達成する場合には、デコーダできわめて大きなバッファが必要になるだろう。

【0085】エンコードシステムによって手動編集と呼ばれることが可能になるが、手動編集とは、映画の中で場面を短縮または延長したり、異なった場面を置換することではなく、画質を変更することである。画質は経時的に変更することができ、本項に説明するように時間的編集または時間的手動編集と呼ばれる。別の項では、フレームのあるセクションからフレームの別のセクションにビットをシフトすることによって、データのフレーム内での画質を変更する、空間編集または空間手動編集と呼ばれることについて説明する。

【0086】時間的編集、つまり編集セグメントのさまざまなセクションの間でのビットを再割当てで、以下に示す3つの主要な制約を満たす必要がある。

【0087】1. 編集される映画/ビデオの総容量が割り当てられた容量を上回らないこと。

【0088】2. 編集セグメント内にバッファアンダフローがないこと。および

3. 時間的編集によって生じさせられる編集セグメントの外側にバッファアンダフローがないこと。

【0089】制約1は、結果として生じる編集済みのセグメントが最初にエンコードされたセグメントと同じビット数を消費するように編集セグメントをエンコードし直すことにより処理される。変種セグメントが検討される期間全体であることを注記すべきである。編集セグメント内の異なる期間には、ユーザにより品質が指定され、編集セグメントのこれらの期間は単にセクション、または期間と呼ばれる。

【0090】アンダフロー状態が存在するかどうかを突き止めるためには、デコーダのモデルを使用して目標ビットを確認することが必要である。エンコーダのVBR動作では、ビデオバッファ検査機構 (VBR) モデルの修正されたバージョンが使用される。このモデルは、VBRの簡略修正モデルと呼ばれ、デジタル記憶装置媒体 (DSM) に使用できる。VBRモード用のVBVは既知でありMPEG-2規格で定義され、ISO/IEC 13818-2の第C.3.2項の付録Cに特に記載される。

【0091】VBR用VBVモデルでは、デコーダのバッファはいっぱいのままであるが、アンダフローしないように制限されている。さらに、バッファは一定のレートRmaxでバッファがいっぱいになるまで充填し、それからバッファが空になり始めるまで、ビットはバッファに伝送されない。各フレーム間隔ごとに、あるフレームを復号化するために必要とされるビットがバッファから削除され

る。これを念頭に入れておくと、VBRモードでは、ピクチャの列に割り当てるビットが多すぎると、バッファアンダフローが発生することが分かる。ピクチャのビットがバッファから削除されるに従い、バッファは空になり、高画質ピクチャのビットの大多数がバッファを充電できるレートより早いレートで削除される場合、バッファがアンダフローする。アンダフローを防止するために、ピクチャをエンコードするために使用されるビット数が低減される。ピクチャのビット数を低減すると、充填レートは同じに保たれるが、ビットをバッファから削除するレートが低減される。VBV占有がいっぱいである場合、VBVバッファへのビットストリーム伝送は即座に停止されるため、バッファオーバフローはVBRに関して発生できない。つまり、バッファオーバフローではなく、単にバッファがいっぱいな状況が存在する。ISO/IEC 13818-2、付録C、第C.3.2に説明されるように、VBRの簡略修正VBVモデルは以下の通りである。

【0092】1. sequence_headerでのbit_rate_fieldはmax_bit_rateを表す。

【0093】2. VBVは当初空である。VBVバッファ（つまりVBVバッファファイル）をmax_bit_rateのレートでいっぱいに満たしてから、復号化プロセスが開始する。

$$R = (\text{総割当てビット} / \text{フレーム数}) * \text{Frame_Rate} \quad (2)$$

総割当てビットはビットレートに正比例しているので、割当てビット数を参照することは、異なる単位を使用する場合を除き、ビットレートと同じ意味を伝える。保証

$$T_i = (R_i - R_{\min}) * N_i / \text{Frame_Rate}$$

(3)

この場合、 N_i はセクション「i」のフレーム数で、 R_i は同じセクションの有効レートである。保証された最小値は、セクションに最小品質レベルを保証するために必

$$T_j = (R_{\max} - R_j) * N_j / \text{Frame_Rate}$$

(4)

この場合、 R_{\max} は最大レート、つまりバッファがいっぱいでないときに充填するレートであり、 N_j はセクションjのフレーム数である。

【0098】期間で画質を変更すること（時間的編集）についての説明は、今度は図12およびそこに参照される図に関して行う。開始後、ステップ281では、第II項および第III項で前述したようにユーザの介入なく自動ビデオエンコードを実行する。自動エンコードの場合、作成されるデータストリームの最大レートおよび最小レートがある。最大ビットレートは、ビデオデコーダの機能により決定され、例えば R_{\max} は毎秒Mbitsである。最小ビットレートは、希望される最小画質に応じて、ユーザが指定できる。典型的な最小ビットレート R_{\min} は毎秒3 Mbitsである。ビデオの部分が低いビットレートとなり、ビデオの部分が低いビデオレートとなることができるよう、平均エンコードレートとは異なり、平均エン

コードレートは、MPEG-2国際規格に定義されるCBRのVBVモデルのようにただちに削除、エンコードされる。

【0094】3. ビットストリームは、VBVバッファがいっぱいになるまでmax_bit_rateのレートでVBVバッファの中に格納される。VBVバッファ占有がいっぱいの場合、VBVへのビットストリーム転送は即座に停止される。次のピクチャのデータが削除されてから、ビットストリーム伝送は即座に開始する。

【0095】4. VBVバッファはアンダフローしないものとする。

【0096】5. 各vbv_delayフィールドがFFFFに設定される。

【0097】期間でビデオの品質を変更するには、ビデオのフレームにビットを追加またはビデオのフレームからビットを削除する必要がある。ここで、ビデオのセクションに追加またはビデオのセクションから削除されるビットの数についての説明を行う。ビデオのセクションに追加またはビデオのセクションから削除できるビットの数は、検討されているビデオのセクションの有効ビットレートに依存する。有効ビットレート R_i は、以下のよう

に計算される。有効ビットレート R_i は、以下のよう

に計算される。有効ビットレート R_i は、以下のよう

要とされる。同様に、指定されたセクション「j」に追加できる最大ビット数は、以下の通り計算される。

コードレートをやや下回る R_{\min} を設定するのが望ましい。 R_{\min} がビデオの平均エンコードレートをわずかに下回って設定されると、結果として生じるエンコードビデオはかなり一定した画質となる。通常、 R_{\max} =毎秒8 Mbitsで R_{\min} が毎秒約3 Mbitsの場合、ビデオの平均値は毎秒約3 1/2から5 Mbitsとなるはずである。

【0099】移動ビデオエンコードがステップ281で実行された後、レート量子化関数を求める必要がある。本発明のプロセスでは、希望の品質を獲得するのに必要なビット数が求められるが、再エンコード時には、希望のビットレートを結果的に生じさせるために量子化値を設定しなければならない。レート量子化関数は、それぞれ参照により本明細書に取り入れられる記事、画像処理に関するIEEE会報、第3巻、第5号、1994年9月、533-545ページのラムチャンドラン(Ramchandran)その他による「マルチ解像度およびMPEGビデオディスクローギャに適

用される依存量子化のためのビット割当て」、IEEE ICA SSP、1993年、V381-V384、ラムチャンドランその他による「MPEGビデオ符号器に適用される依存量子化のためのビット割当て」、および1994年にプリンストンビデオ会議で提示されたレイニング(Reininger)による「レート歪み曲線を使用する多重化されたMPEGビデオ用の共同マルチチャネルビットレート制御」に説明されるような従来の方法で計算できる。

【0100】次に、ステップ284では、選択された期間のユーザ品質設定値を入力する。ユーザは、究極的にはピクチャの量子化を制御するのに使用される異なったセクションの優先順位を入力する。量子化は、損失のあるMPEGエンコードの解像度とも呼ばれる画質を制御する。図13には、期間でビデオの品質を変更するために情報を収集する目的で使用されるユーザインタフェースを説明する。当初、ユーザは、品質を変更することになる編集セグメント全体の期間を入力する。この応報の入力は、図13のユーザインタフェースには図示されていないが、図13のユーザインタフェースが表示される前に入力される。図13に示される例の場合、変更対象の編集セグメントの期間が、例えば時間0分から時間5分までとなることがある。ユーザが編集セグメントの総期間を記述した後、編集セグメント内のさらに短い期間に、優先順位Pを含む-5から+5までの範囲のさまざまな優先順位が指定され、この期間の品質が保護され、改変してはならないことを示している。優先順位「なし」は、Pという優先順位と同じ意味である。エンコードシステムは前記入力スキームに限定されず、ユーザの希望の優先順位を入力するために、それ以外の範囲または記号を使用できるのと言うまでもない。

【0101】優先順位を指定するのに加えて、ユーザは最小ビットレートも入力する。このビットレートは、ステップ281の自動ビデオエンコードの間にビデオをエンコードするために使用される最小ビットレートと異なっている場合がある。ユーザインタフェースのもう1つの特徴とは、それを使用すると、ユーザが「バックグラウンド優先順位」を設定できるという点である。これは、ユーザが優先順位を指定したのではない選択済みの期間内の期間にこのバックグラウンド優先順位が設定されることを意味する。例えば、図13では、バックグラウンド優先順位に「1」が指定されるとすると、未指定の期間00:03:20:01から00:03:59:99には、ユーザが手動でこの特定のセクションに優先順位を指定しなくても、「1」という優先順位が自動的に指定される。

【0102】図13に描かれるインタフェースの底部では、編集セグメント内で新しい期間を作成するために「挿入」ブロックが使用され、以前に作成された期間を変えるために「リコール」ブロックが使用され、期間に加えられる変更を保存するために「更新」が使用され、期間を削除するために「削除」が使用される。光ディス

クのようなエンコードされたビデオ用の究極のデジタル記憶装置媒体は、自動的にエンコードされたデータがディスクに書き込まれる場合には、理想的には容量まで満たされる制限された記憶領域容量を備える。したがって、画質を向上させ、ある特定の期間のビットレートを引き上げる場合はつねに、品質が向上した期間に必要な余分な記憶容量を供給するために、品質が低下したビデオのセクションが存在しなければならない。したがって、本発明の実施例では、品質を改善するのに必要なビットはビデオの別のセクションから取らなければならないため、別の期間に違う画質を割り当てないで、1つの期間だけに高い画質を割り当てることは不可能である。したがって、編集されたビデオのセグメントには、バックグラウンドセクションとは異なり、時間的な手動編集が適切に動作するために「保護つき」のセクションとマークされていない少なくとも1つのセクションがなければならない。さらに、ユーザが優先順位を設定した結果、同じ画質およびビットレートが、それらが選択されるたびに生じることはないが、編集セグメント内のピクチャのビットの修正された数を求めるためのプロセスを理解すると明らかになるように、ある特定のセクションの品質は、他のセクションの割り当てられた優先順位および他のセクションの長さにも依存する。

【0103】本発明によるエンコードシステムにより活用される特徴とは、視聴者が、自動エンコードプロセス中にまたは後日、エンコードされたビデオをレビューすることができるという点である。エンコードされたビデオが視聴者に表示されるに従って、視聴者は、ビデオを停止したり、関心のあるタイムコードを書き残したりしなくても、ビデオの関心のあるポイントや領域を記憶するために、リターンキーを押したり、画面に表示される「ボタン」の上でグラフィックユーザインタフェースを「クリック」するなどの機能を実行できる。これらの関心のあるポイントは、後日、ユーザが呼び出し、映画全体をレビューしなくても綿密に検討することができる。これは、ユーザが変更を希望するエンコードビデオ内のポイントを決定する便利で、効率的な方法となる。この機能は、関心のあるポイントや領域が発生した時点での位置に対応して、タイムコードを記憶することによって達成される。

【0104】ユーザがステップ284で希望の品質設定値を入力してから、ステップ286でユーザの入力に基づいてフレームごとに修正されたビットが計算される。このプロセスは、図14-16Bに関して詳細に記述される。ビデオの編集セグメント内で時間のビットレートを変更するための一般的な手順は、図14のフローチャートに説明される。ユーザの入力に基づいてビットレートを計算するために、ステップ302では、-5と+5の間の範囲内から10と0の間の範囲内にあったユーザの入力をマッピングする。-5という初期入力、10というマッピングされた設

定値に対応し、最小品質であり、+5という入力値は0というマッピングされた設定値に対応し、最大画質である。同様に、0という入力された品質選択は5というマッピングされた設定値に対応するなどである。ステップ302で実行されるマッピングの目的とは、割り当てられたビット数がマッピングされた設定値により乗算される量-1のべきまで引き上げられる $e(2.71828\dots)$ に比例することである。この場合、優先順位の設定値がさらに小さい場合のビット数が、優先順位が高い場合より下回るように、さらに $e-10$ は $e0$ を下回る。目標ビットを優先順位-4から+4に対応して求めるプロセスは、図16Aと図16Bに詳細に説明される。

【0105】ステップ304では、「保護つき」のセクションがあるかどうか判断される。保護つきセクションとは、画質を以前の自動エンコードから変更してはならないセクションのことである。保護つきセクションがある場合は、フローはステップ306に進み、そこで保護つきセクションの再エンコードに割り当てられるビット数が最初の自動エンコードの結果生じる以前のビット数に等しくなるように設定される。このようにして、保護つきセクションのピクチャ品質は変化しない。

【0106】ステップ308では、最大品質または最小品質に設定されるセクションがあるかどうか判断される。最大品質とは、ユーザ入力+5に対応し、最小品質とはユーザ入力-5に対応する。最大画質および最小画質の場合、依然としてビデオバッファの制約を満たし、エンコードされたビデオを記憶する媒体の記憶容量を上回らずに、それぞれ最高品質のビデオと最低品質のビデオが設定されたエンコードされたビデオを入手するのが望ましい。最大品質または最小品質に設定されるセクションがある場合は、フローは、ステップ310に進み、そこで最大ケースまたは最小ケースの修正ビット数が計算される。このプロセスは図15に詳しく示される。

【0107】図15では、最大ケースおよび最小ケースに割り当てられたビット数を求めるプロセスが、マッピングされた優先順位の10と0にそれぞれ割り当てられたビデオのセクションに最小レートおよび最大レートを割り当てることによってステップ320で開始する。図14のステップ302では、このマッピングおよび時間的編集の記述の残りの部分が実行され、それ以外の明確な指示がない限り、マッピングされた優先順位が参照される。以下に記述されるように、最大データレートはデコーダの能力により制限を受け、毎秒1 Mb/sとなり、最小データレートは、示しているように図13でユーザインタフェースによりブロック図され、通常は毎秒3 Mb/sとなる自動エンコードモードの間に使用される最小ビットレートを下回ることもあれば、上回ることもある。

【0108】それから、ステップ322では、マッピングされた優先順位1から9が設定されたセクションがあるかどうか判断される。つまり、最大品質または最小品質が

割り当てられていることに加え、ユーザが割り当てるそれ以外の中間品質が存在するのである。それ以外のマッピングされた優先順位がある場合には、フローはステップ324に進み、1と9の間のマッピングされた優先順位が設定されるすべてのセクションの平均ビットレートが計算される。この平均が許容範囲（最大ビットレートと最小ビットレートの間）にない場合、フローはステップ328に進む。ステップ328では、最大レート、つまり最高の画質を実現するレートは、平均が最小レートを下回る場合に低減される。最大レートを引き下げることにより、非最大状況および非最小状況（マッピング優先順位1-9）に使用できるビット数が多くなり、それにより平均が希望の範囲内または希望の範囲近くとなる。代わりに、ステップ328では、平均が最大レートを上回る場合、最小データレートは引き上げられ、それにより非最大状況および非最小状況（例えば、優先順位1-9）のビット数が引き下げられる。ステップ328の後、フローは平均が範囲内になるまでステップ324、326、および329が繰り返される324に戻る。さらに、最大ビットレートおよび最小ビットレートは、「保護つき」セクションが品質の調整の前と同じビットを受け取ることを確実にするために調整される。

【0109】いったんステップ326で平均が範囲内にあると判断すると、フローはステップ330に進み、それぞれマッピングされた優先順位10と0が設定されるセクションに計算された最小レートおよび最大レートが割り当てられる。それから、ステップ332で、残りのビットは、保護つきではなく、最大状況や最小状況でもないセクションに分配される。マッピングされた範囲1-9が設定されるビデオのセクションおよびデータレートの変更方法については、図16Aおよび図16Bに関して後で説明する。

【0110】ステップ322でマッピングされた優先順位1-9が設定されるセクションが存在しない、つまり最大セクション、最小セクションまたは保護つきセクションだけが存在すると判断した場合は、ステップ340で、ビットの最終割り当てが自動エンコードによるビットの初期割り当てと等しくなるまで、必要に応じて最小レートおよび最大レートを調整する。ステップ340を実行する2通りの代替方法がある。第1に、最大ビットレート R_{max} は、例えば毎秒8 Mb/sで固定されている。総合的な編集セグメントの合計ビットは理解されているため、最小データレートである R_{min} は、自動エンコードの間に獲得されると同じビット総数を達成するために調整されるだけである。 R_{min} を総ビットの以前の数を達成するために調整することができない場合には、自動エンコードの間に獲得されるビット総数を達成するために、最大データレート R_{max} が調整される。

【0111】代替策として、ステップ340では最大品質セクションに優先順位1および最小品質セクションに優

先順位9が指定され、編集セグメントの期間の新しいビットレートを求めるために、図16Aおよび図16Bにブロック図で示されるビット割当てアルゴリズムが実行される。それから、ステップ342で図16Aおよび図16Bに従って計算されたデータレートが、以下に説明するように、最低指定品質および最高指定品質が設定されるセクションに割り当てられ、図15のプロセスが終了する。

【0112】ステップ308およびステップ310が、必要な場合に実行された後で、ステップ312で、すべてのセクションにマッピングされた優先順位1-9が設定されるかどうか判断する。これが当てはまると、フローはステップ314に進み、図16Aおよび図16Bに詳説するように、マッピングされた優先順位の修正ビット数が計算される。

【0113】図16Aの最初のステップでは、使用される共通ビットプールのサイズを決定する。当初、保護つきではないマッピング優先順位1-9が設定されるすべてのセクションで、一定数のビットが削除されている。再分配深度と呼ばれる係数 k_1 は、このために使用され、例えば、0.4に設定され、以前に（つまり自動エンコード実行から）割り当てられたビットの40%が、保護つきでもなく、最大品質や最小品質によりマークされてもいないあらゆるセクションから削除されることになることを意味する。

【0114】次に、ステップ362では、ユーザ選択品質がどの程度品質に影響を及ぼすかを判断する。動的範囲係数と呼ばれる変数 k_3 が設定され、優先順位から生じることになる品質変更を決定する。 k_3 がどのように設定されるかに応じて、例えば優先順位2は、品質での相対的に小さな向上、または品質での比較的に大きな向上のようなさまざまな品質の変更を課す場合がある。通常、動的範囲係数 k_3 は、1.0に設定される。

【0115】ユーザ選択優先順位をビットの再分配で使用するには、形式 $\exp(-\text{優先順位}/k_3)$ の指数的な関係が使用される。指数内の動的範囲係数 k_3 が、指数曲線の形状を決定する。曲線は、動的範囲係数の値が大きくなるほど急勾配となり、実質上、変化する優先順位が設定されたセクションの間で割当てし直したビットでの不均衡が強まる。この負の指数は、マッピングされた優先順位の絶対値が増加するに従い（さらに低い優先順位セクション）、そのセクションに割り当てられるビットは減少する、つまりマッピング優先順位1は最高の優先順位で、9まで数を引き上げると設定される優先順位が低くなることを意味する。

【0116】ステップ364では、ユーザ選択優先順位が正規化される。つまり、ユーザが選択した優先順位が、各セクションの相対優先順位を決定するために使用される。まず、 E_1 が計算される。

$$E_1 = \exp(-p_1 / k_3)$$

(5)

この場合、 p_1 とはマッピングされたユーザ選択優先順位

で、 i は、保護つきまたは最大品質や最小品質が設定される期間を除く、編集セグメント内のセクション期間である。次に、優先順位製機化係数 k_2 が、以下のように計算される。

【0117】

$$k_2 = \sum_{i=1}^N E_1$$

(6)

この場合、 N はセクション数である。優先順位正規化係数は、優先順位に従って割当てし直されたすべてのビットの合計が、つねに、ビットの使用可能な共通プールに等しくなることを確実にするために必要とされる。

【0118】したがって、各セクションの相対優先順位は、以下のように計算される。

【0119】

$$R_i = E_1 / k_2$$

(7)

R_i は i 番目のセクションの相対優先順位の端数であるので、すべての R_i の求和は単一となる。

【0120】ステップ366でセクションのサイズが正規化される。割り当てられた優先順位が設定される期間のサイズは異なり、ビットはさまざまなセクションのサイズに比例して分配されなければならないので、これが必要である。第1に、編集済みフレーム総数での各期間内のフレーム数の端数 T_i （保護つきではない、非最大、または非最小優先順位が設定されるフレーム）は、以下に従い求められる。

【0121】 $T_i = N_i / N$ (8)

この場合、 N_i は i 番目の期間内のフレーム数であり、 N は編集セグメントのフレーム総数である。つまり、改変され、保護されておらず、最大でも最小でもない優先順位ではない優先順位が指定されるすべてのフレームである。したがって、サイズ正規化係数 k_3 は、以下の等式に従い計算される。

【0122】

$$k_4 = \sum_{i=1}^N (T_i * R_i)$$

(9)

この場合、 N はセクション数である。図16Aのステップ366から、フローは図16Bのステップ368に進む。

【0123】図16Bでは、ステップ368で、以下の等式に従い、ビットの共通プールが求められる。

【0124】

$$C = \sum_{i=1}^N (k_i * B_i)$$

(10)

この場合、 B_i は i 番目のセクションの自動エンコード手

順の間に使用されたビット数に図15のステップ332で分配されたあらゆるビットを加算したビット総数である。求和は「N」セクション上である。

【0125】それから、ステップ370で、以下の等式に従い各セクションに割り当てられるビットのターゲット数を求める。

$$F_i = C * T_i * R_i / k_4 + (1 - k_1) * B_i \quad (11)$$

各セクションのビット総数Fiは、ビットの共通プールのパーセンテージにビットの共通プールに与えられたビット分低減されたビット(Bi)の元の数を加算したものに等しい。ビットの共通プールからのビットのパーセンテージは、各セクションのフレーム総数の端数(Ti)、セクションの相対優先順位(Ri)、およびセクションのそれぞれに総計されるフレームおその対応する端数で乗算されるセクションの優先順位に依存するサイズ正規化係数(k3)に基づいている。

【0126】ステップ370で計算されたセクションごとに割り当てられたビットのターゲット数が最大データレートと最小データレートの間でない場合がある。これが当てはまると、ステップ372で、再分配深度k1が0.01低減され、フローは、ステップ368のビットの共通プールを計算し直し、ステップ370のビットのターゲット数を計算し直すために戻る。ステップ368、370、372および374のループが、ステップ370で計算されたビットのターゲット数が最大データレートと最小データレートの間になるまで実行される。

【0127】それから、ステップ376では、ビット位取り因数Siおよびイントラフレームおよび非イントラフレームが、希望の品質を獲得するために異なるビット数が必要とする。まず、ビット位取り因数Siが、以下の等式に従い計算される。

$$S_i = F_i / B_i \quad (12)$$

次に、イントラピクチャおよび非イントラピクチャの位取り因数が計算される。イントラ位取り因数Sliは、以下の等式に従い計算される。

$$S_{li} = S_i - (I_i * S_i') \quad (13)$$

大きなイントラフレームにより、バッファアンダフローが発生する可能性があり、イントラ因数Ifは、イントラピクチャのビット割当てでの変動を、If=0.8のときに20%に制限することによって安全率を示す（つまり、Sliは範囲0.8<Sli<1.2に制限される）。Si'-Si-1.0であることを注意する。非イントラビット位取り因数Sniは、以下のように計算される。

$$S_{ni} = S_i - (I_i * S_i' * B_{li} / B_{ni}) \quad (14)$$

この場合、BliおよびBniは、ピクチャタイプによってi番目のセクションで合計される総イントラビットと非イ

ントラビットである。

【0128】最後に、フレームごとに使用されるビット数は、適切な食らい取り因数（イントラ位取り因数または非イントラ位取り因数のどちらか）で乗算される元のビット数を使用して求められる。それから、図16B（および図14）のプロセスが終了する。

【0129】この時点で、各フレームの修正されたビット数が計算され、図12のステップ286が完了する。ただし、フレームの変更済みビット数によりバッファアンダフローが生じたり、編集セグメントの立ち下がり端でのビット数の占有が、編集セグメントの外のそれ以降のフレームがアンダフローするように変化する可能性がある。これらの状況は、必要ならば、図17に詳説される図12のステップ288で調べ、訂正される。

【0130】図17は、バッファアンダフローがないか確認、訂正し、編集セグメントの立ち下がり縁でのビット再割当てのための問題を確認、訂正するプロセスを説明する。図17では、ステップ390で編集セグメントのフレームのそれぞれのビット割当てをスキャンする。それから、ステップ392で、前述のVBVモデルを使用する新しいビットの分配のアンダフローがないか確認する。アンダフロー問題が存在し(u_flowリスト)、各アンダフローを排除するのに必要なビット数が求められ、u_valに記憶されるポイントのリストが作成される。アンダフローが、多すぎるビットによって表されるピクチャのために発生することを頭に入れておくことが重要である。多数のビットで表される高品質のピクチャがデコーダによってバッファから読み出される場合は、ピクチャが復号化されるときにバッファから多数のビットが削除され、これらのビットは十分に速く再び満たすことはできない。したがって、アンダフローを排除するためには、ビットはピクチャ空削除される。さらに、ステップ392では、バッファがいっぱいであるポイントのリストが作成され、これらのポイントはo_flowに記憶される。前述したように、バッファがいっぱいになると、データはバッファに伝送されなくなるので、バッファがいっぱいであっても問題はないことに注意する。

【0131】ステップ394では、編集セグメントの立ち下がり縁でのバッファ占有を調べる。動作のVBRモードでは、編集済みビデオセグメントの立ち下がり縁でのバッファ占有が、編集済みセグメントがまったく同じビット数で置き換えられていても、修正された実行のバッファ占有とは異なる可能性がある。これは、VBR用のVBVモデルの特異性から生じる。したがって、バッファ占有の制約を調べるのが重要である。編集セグメントの立ち下がり縁でのバッファ占有がビットの再分配の前とまったく同じか、あるいはさらに高い場合に、問題は存在しないことを示すのはかなり簡単である。バッファ占有が以前とまったく同じ場合には、それ以降のフレームの占有も、VBV違反がないかすでに確認済みの以前の実行とま

まったく同じままとする。占有が以前より高い場合、バッファがそれ以降のフレームの間にいっぱいになる可能性があり、それは、バッファがいっぱいな場合、データはバッファに送られなくなり、アンダフローは発生しないので完全に許容できる状態である。しかし、以前の実行と比較して、バッファ占有が立ち下がり縁でさらに低い場合に、第3のケースが生じる。これは、それ以降のフレームでバッファアンダフローを作成する可能性があり、密接に調べる必要がある。編集セグメントの最後のフレームを、バッファセグメントの外側でのフレームのバッファアンダフローを防ぐために処理できる2通りの方法がある。第1のアプローチは、2つの内で簡単な方である。編集セグメントの最後のフレームのビット数により、(ログファイルに記憶される)元のエンコードより大きいまたは元のエンコードに等しいバッファ占有が生じる場合、アンダフロー問題は存在せず、したがって処置を講じる必要はない。編集セグメントの最後のフレームのバッファ内のビット数が、最初にエンコードされたセグメントのバッファ内のビット数を下回る場合、ビットのこの低減された数により、編集セグメントの後のフレームの内容に応じて、編集セグメントの外側のフレームのアンダフローが生じる場合もあれば、生じない場合もある。アンダフローが発生しないように防ぐのに必要となるビット数を求める代わりに、編集セグメントの最後のセグメントのバッファの占有が、元のエンコードの場合と同じになるように調整されるだけである。安全性の目的から、編集セグメントの最後のバッファ占有または十分さは、アンダフローの可能性がないことを保証するために元のエンコードの占有の10%上まで引き上げることができる。編集セグメントの最後のフレームのバッファ占有を元のエンコードのレベルになるように引き上げる必要がある場合は、終了フレームのビット数をステップ398で実行したように低減する必要がある。最後のフレームはアンダフローポイントとして処理され、u_flowリストに格納され、バッファレベルを元のエンコードのレベルまで引き上げるために最後のフレームから削除する必要があるビット数がu_valに付加される。

【0132】編集セグメントの再エンコードにより生じるアンダフローを処理するための第2のアプローチは、編集フレームの外側でバッファアンダフロー状況を引き起こさない編集セグメントの最後のフレームのバッファの最小可能レベルを概算する反復プロセスである。これは、元のエンコードのバッファ占有と編集セグメントの最後のフレームの再エンコードのバッファ占有の間の最大差異を概算するプロセスにより実行される。

【0133】図18にブロック図で示されるフローチャートを見ると、元のエンコードバッファ占有にはB0が割り当てられ、元のエンコードのログファイルに記憶されていた。i=1...nの場合に編集セグメントの境界を超えたフレーム「i」でのバッファ占有はBiである。この場

合、「n」はエンコードされる映画内の最後のフレームに相当する。再エンコードのバッファ占有が、連続するフレームのどれかでアンダフロー状況を引き起こさずに元のエンコードからどの程度低くなることができるかは、再エンコードのバッファレベルを、少なくとも、それが元のエンコードのためであったポイントに設定するだけの第1アプローチとは対照的に決定されなければならない。この条件を満たす境界での新しいさらに低いバッファ占有をB0'にする。これは、編集セグメント境界での最小許容バッファ占有である。以前の実行のバッファ占有より低いバッファ占有を設定する可能性は、差異が、編集セグメント後のもっと遅いフレーム間隔で吸収されるため、アンダフローが発生しないという事実から生じる。

【0134】図18のステップ420で開始し、編集セグメントの最後のフレームの再エンコードされたビデオのバッファ占有B0'が、将来アンダフローを引き起こすことなくゼロであると想定する。これは、バッファ占有が、アンダフローを防ぐためにはより高くなければならないが、バッファ占有はゼロを下回ってはならない可能性があるため、B0'に関してもっとも自由な概算である。しかし、図18のプロセスは、再エンコードされた編集セグメントの最後のフレームのゼロの占有によりアンダフローが生じるかどうか、および生じる場合、アンダフローを引き起こさない値B0'が見つかるまで、編集セグメントの最後でバッファ占有を増加させるかどうかを確認するためにチェックする。

【0135】図18のプロセスでは、元のエンコードと再エンコードの間の概算バッファ占有差異「X」を評価する。

【0136】

$$X' = B_0 - B_0' \quad (15)$$

差異「X」は、「X」がチェックされ、必要に応じてアンダフローのすべての可能性が排除されるまで調整されていることが保証されるまで、最終的に決定できないので概算と呼ばれる。図18の手順は、境界フレーム直後のフレーム、つまりi=1で開始して、一度に1フレームずつ進む。B0'がゼロに等しい第1フレームの場合、概算差異は、ステップ422で計算されるように、元のエンコードの第1フレームのバッファ占有の値であるのにすぎない。

【0137】

$$X' = B_0 \quad (16)$$

次に、ステップ424では、境界フレームの右側にあるフレームにより吸収されるビット数を求める。吸収されるビット数とは、元のエンコードの編集セグメント後のフレームによりバッファの中に格納される場合もあるが、実際には、バッファがいっぱいであるためにさらにビッ

トを受け入れることができなかったために元のエンコードの間バッファに格納されなかったビットを指す場合に使われる言い回しである。量 Δti は、バッファがその間いっぱいであったが、 B_0' の値が引き下げられたために現在では追加ビットを受け入れることができる編集セグメントの後の元のエンコードの i 番目のフレームの期間を表す。バッファはフレーム全体でいっぱいとなることはできず、ゼロより大きい、ゼロに等しいため、期間 Δti はフレームの期間を下回らなければならない。元のエンコードの間にバッファが一度もいっぱいではなかった場合には、期間 Δti はゼロである。したがって、

$$1/P > \Delta ti \geq 0$$

(17)

この場合、 P' は毎秒フレーム単位で測定され、通常、毎秒30フレームであるピクチャレートである。バッファがいっぱいである結果、編集セグメントの右側で吸収されるビット総数は、以下のように計算される。

【0138】

$$a_i = \sum_{k=1}^i \Delta t_k R_{\max}$$

(18)

この場合、 R_{\max} はバッファの充填レートである。

【0139】ステップ426では、編集セグメントの最後のフレームのバッファ占有レベルを求めるための手順が終了できるかどうかを判断する。 i 番目のフレームによって吸収されるビットの合計が X' を上回る、つまり $a_i > X'$ である場合は、バッファアンダフローが発生しないで、ステップ428が実行されるように、バッファはもっと遅いフレームで満たされることができ、手順は停止される。代わりに、最後のフレーム n に到達すると($i=n$)、ビデオの最後まで X' の概算値によりアンダフロー問題が生じることはない。

【0140】ステップ426に対する回答が「ノー」の場合、ステップ430で吸収されずに残るビット数 b_i が計算される。

【0141】

$$b_i = X' - a_i$$

(19)

それから、ステップ432では、 $b_i > B_i$ であるかを判断する。この場合、 B_i とは編集セグメントの後の i 番目のフレームのバッファ占有である。 $b_i > B_i$ の場合、現在のバッファ占有「 B_i 」が許容するより多くのビットが吸収されないで残っているため、差異の概算値、 X' は大きすぎて、アンダフローを生じさせる。したがって、吸収されずに残るビットの間では、「 B_i 」ビットだけがアンダフローを引き起こさずに吸収できる。次に、差異の概算 X' はステップ434で改訂されなければならない。ステップ434では、 X' の改訂された概算は、吸収されずに残るビットに吸収されたビットを加算したものと計算され

る。

【0142】

$$X' = B_i + a_i \quad (20)$$

ステップ432で、 $B_i < B_i$ であると判断する場合、 i 番目のフレームでのバッファ占有は吸収されないで残るビットより大きいので、アンダフローの問題はなく、 X' の概算を変更する必要はなく、フローはステップ436に進む。ステップ436では、次のフレームが調べられ(i は1で増分される)、ステップ424、426、430、432、434、および436のループが、ステップ426での判断に肯定的な回答が得られるまで繰り返される。

【0143】ステップ426に肯定的な回答が得られると、概算 X' は受け入れられるため、ステップ428では、編集セグメントの境界での最小許容バッファ占有 B_0' は以下の通りである。

$$B_0' = B_0 - X'$$

(21)

編集セグメントの再エンコードされた最後のフレームの実際のバッファ占有が B_0' を下回る場合、アンダフロー状況がビットの再割当てのために発生する。この場合、最後のフレームは u_flow リスト内のアンダフローポイントとしてマークされ、アンダフロー状況(変更済みビデオのビット数と結果的に最小許容バッファ占有を生じさせるビット数 B_0' の間の差異)を防ぐための編集セグメントの最後のフレームから削除する必要があるビット数は、ステップ392について記述されたのと同様に u_val の中に格納される。

【0144】本発明のエンコードシステムの場合、ビデオの中のすべてのポイントのバッファ占有を知ることが望ましく、この情報はログファイルに記憶されることに注記すべきである。しかし、編集セグメントの最後のフレームでのバッファ占有を引き下げる場合には、編集セグメント後のフレームのバッファ占有が変更され、ログファイル内のバッファ占有情報が不正確になり、そのため編集セグメント後のフレームのバッファ占有を計算し直す是正処置を講じる必要があるだろう。

【0145】図17のプロセスに戻って参照すると、ステップ396では、 u_flow リスト内にエントリがあるかどうかを判断する。ない場合は、アンダフロー問題は存在せず、図17のプロセスは終了する。 u_flow リストにエントリがある場合は、フローはステップ398に進み、 u_flow リスト内の各位置の左側からもっとも近い u_flow ポイントまでのセクションが保護つきとマークされ、 u_val 内の対応するビット数に等しいビットがこれらのセクションから削除される。バッファアンダフローは1つ以上のフレーム内の多すぎるビットにより引き起こされるため、フレーム内のビット数を低減すると、アンダフロー問題が解決される。アンダフローを防止するために削除

されるビットが原因で画質に著しい低下が生じないようにするために、ビットはアンダフローポイントだけから削除されるのではなく、バッファがいっぱいであった最後のポイントまで戻ってすべてのフレームから均一に削除される。これが、アンダフローを防止しながらも最高品質のビデオを獲得する方法となる。

【0146】その後、ステップ398で削除されるこれらのビットは、ステップ400で保護つきとマークされないセクションに均一に分配され、図17のプロセスがもう一度開始する。ステップ400でビットを分配しても、アンダフローが問題ではなくなるという保証にはならないため、図17のプロセス全体を繰り返し使用して、アンダフロー状況がないか再度確認する必要がある。

【0147】図17のプロセスが完了した後、図12のステップ288は完了し、図12のステップ290が実行される。ステップ290では、ステップ282で求められるレート-量子化関数を使用してフレームごとの平均量子化値を求める。希望されるビットの数は、以前ステップ286で計算され、ステップ288でアンダフロー問題が存在するかどうかを確認するためにチェックされたが、エンコードされたデータに希望のビット数が設定されるようにデータを再エンコードするためには、量子化値を求めなければならない。これは、レート-量子化関数を求め、レートを捜し出すことで量子化値を求めるだけで、決定される。

【0148】ピクチャ単位で量子化レベルを求める場合は、レート-量子化関数(RQF)は、ステップ282で求められた。この関数を概算するには、ピクチャごとに2つのデータポイントが必要とされる。RQFは以下の通りである。

【0149】

$$R = X * Q^{(-R)} + H \quad (22)$$

この場合、Rは1つのピクチャをコーディングするために使用されるビット数、Qは量子化スケール、X、g、およびHはピクチャ単位のコーディング特性パラメータである。Hは、Hがヘッダにより消費されるビット、運動ベクトラ、DC係数などとなるように、量子化スケールとは独立したビットを表す。

【0150】エンコーダの最初の実行は、ビデオシーケンス内のコーディングされたピクチャごとに2つのデータポイントを作成するために使用できる。これは、ピクチャ内のマクロブロックごとに使用される1組の交互の量子化スケールにより実行できる。これらは、レート-量子化モデルパラメータ「X」と「g」を計算するために処理される。これらの2つのパラメータは、Hの値とともに、コーディングされたビデオのフレームごとに1つ、ファイル内に記憶できる。

【0151】ビデオのある特定なコーディングされたフレームに割り当てられるターゲットビットを作り出す量

子化スケールを求めるために、「R」がターゲットビットを表す前記等式が使用され、(X、g、H)の値が前記のログファイルである可能性がある記憶されたデータファイルから読み取られる。この等式に最適の「Q」の値が希望の量子化スケールである。

【0152】それから、ステップ292でステップ290で求められた量子化値を使用してビデオの変更されたセクションを再エンコードする。再エンコードの間に特別な手順が講じられない限り、既にエンコードされたビデオの再エンコードセグメントによりエンコードによって問題が生じる可能性があることに注記すべきである。エンコード構造および復号化構造を回避するための再エンコードプロセスの詳細は、以下の別個の項に説明する。ビデオは再エンコードされてから、以前エンコードされたビデオのビットストリームに代入され、プロセスは終了する。

【0153】図5に関して記述するように、ユーザは、ビデオが許容できるかどうかを判断するために、ビデオが再エンコードされた後でビデオを表示することができる。ビデオが許容できない場合には、ユーザは新規にエンコードされたビデオを受け入れる必要はなく、既にエンコードされたビデオの使用を続けることができる。ユーザが新規にエンコードされたビデオが望ましいとわかったら、新規にエンコードされたビデオで既にエンコードされたビデオは置き換えられる。それから、ユーザは、図5に説明するように、ビデオの新しいセグメントを、空間的に、または時間的に編集できる。

【0154】図19(A)および図19(B)では、オリジナルビデオエンコードからの、およびユーザが手動品質優先順位を入力した後のビットレートの2つの異なる例を示している。図19(A)および図19(B)では、(メガビットを表す106のスケールでの)ビットレートが、ピクチャのグループ番号に対してプロットされる。ブロック図で示される例では、1つのGOPに15のピクチャが格納され、各GOPは0.5秒ごとに表示される。また、実線は元のエンコードを示し、破線はユーザが手動で画質を調整した後のエンコードを示す。

【0155】図19(A)では、優先順位2、-3、3、0および保護つきが設定される5種類の異なった優先順位領域のあるエンコードを説明する。優先順位3が設定される期間は、再エンコードの後に、元のエンコードに比較してさらに多くのビットを使用する。優先順位はただ単に互いに相関しているだけなので、優先順位2が設定されるセクションのビットレートは、著しく上昇していないが、-3および0のセクションではビットレートは低下している。保護つきセクションには、再エンコードの前に存在したのと類似した再エンコード後のビットレートが設定されなければならない。

【0156】図19(B)は、それぞれユーザが選択した優先順位0、2、0、4、および0が設定された領域のある

第2例である。ここでも、優先順位は互いに相関しているので、優先順位0は、ビットレートが未変更であることを意味するのではなく、単に、優先順位2と4が設定されるそれ以外のセクションに相対する優先順位である。優先順位4にが設定されるセクションに高いビットレートを指定するために、ユーザ選択優先順位0が設定されるこれらのセクションは低減され、優先順位2が設定されるセクションには、エンコードの前後とほぼ同じビットレートが設定される。

【0157】時間的手動編集の処理は、本発明の実施例の図1Aのワークステーション30で発生する。ただし、時間的手動編集はビデオエンコードプロセスの間は発生しないので、ワークステーション10はビデオエンコードプロセスのレートを低下させないで時間的手動編集計算を実行できる。

【0158】本発明が、編集セグメント内でのさまざまな時間セグメントの画質の変更として詳しく記述されることに注記すべきである。言うまでもなく、同じ時間セグメント内にはピクチャのセクションの質の変更を許可するためにこの概念を拡張することは明らかに可能である。例えば、映画の最後の5分間の質を最初の5分間を犠牲にして向上させるために、本発明の教示を適用することができる。以下に説明するように、単独編集セグメントの外側にある期間の品質を変更することが希望される場合、ユーザは、関心のあるセグメントを連続して表示し、チェックし、ビットを割当てし直し、各セクションの立ち下がり縁でのバッファ占有だけではなく、映画の合計ビットに関係する制約が規定の限度内にあることを確認する必要があるだろう。

【0159】時間的手動編集は、編集セグメントの元のエンコードのビット数を、編集セグメントの再エンコードのビット数と同じにしておくこととして記述されてきた。しかし、希望される場合には、再エンコードされた編集セグメントが消費するビットをさらに少なくしたり、十分な記憶容量が存在する場合には、再エンコードされた編集セグメントのビット数が元のエンコードのビット数を上回ることがある。また、ビデオの元のエンコードを、それにより最終デジタル記憶媒体上の使用できるすべての記憶スペースが消費されないように実行することもできる。したがって、再エンコード後のビット総数は、元を上回るまたは下回る任意の量、例えば5%と1%多いまたは少ないビットを含む、0%と20%の間で多いか少ない量を消費できる。

【0160】時間的手動編集の説明は、ビデオエンコード後にビデオを変更することに関して行われてきた。しかし、本明細書に記載される教示は、初期エンコードプロセスがなかったシステムにも適用できる。さらに、共通ビットプールの使用に関する時間的手動編集の教示は、以下に示すビデオフレームの空間手動編集の概念に適用できる。また、空間手動編集の個々のフレームでの

動作様式は、以下に説明するように、上質の時間的編集を達成するためにフレームの列に適用できる。

B. フレームの領域内での品質の変更

図20には、1つ以上のフレーム内で領域の品質を変えるための一般的な手順を説明する。開始後、ステップ450では、システム構成要素および一般的なシステム動作に関する項に説明するように、MPEGフォーマットやそれ以外のフォーマットのような圧縮済みデジタルフォーマットに入力ビデオをエンコードする自動ビデオエンコードを実行する。ステップ452では、ユーザにエンコードされたビデオが表示され、ステップ454では、ユーザは1つ以上のフレーム内の領域の品質の変更を示すコマンドを入力できる。

【0161】図21には、図1Aに示されるビデオディスプレイモニタ61のようなモニタ上で表示されるビデオのフレームがブロック図が示される。図21のビデオのフレームは、木465、人467、鳥468、および2つの雲471があるとして描かれる。

【0162】図21では、フレーム内の領域が、木465を取り囲む領域466、人を取り囲む領域469、および領域470、重複領域469を含めて、ならびに鳥468および人467の頭を含めて限定される。これらの領域は、図1Aに示されるグラフィックポインティングデバイス33を使用して描画された。図21でユーザが設定した品質優先順位には、木を含む領域466の-5、人を含む領域469の+2、および鳥と人の頭を含む領域470の+4が含まれる場合がある。図21内のビデオのフレームのそれ以外の領域には、ユーザ定義優先順位は指定されず、したがって、「無定義」優先順位が指定される。後述するように、「無定義」領域とは、ユーザ定義優先順位が指定されるフレームのエンコードのビット総数を、フレームの元のエンコードの結果生じるビット数に等しくなるように調整する目的で、ユーザ定義領域の量子化レベルを変更した後に最初に修正される領域のことである。ユーザがフレームの領域の品質の変更を希望しない場合、これらの領域は、ユーザによって優先順位ゼロが設定されるとマークされるだろう。

【0163】本発明の実施例では、ビデオのフレーム内で優先順位領域を限定すると、ユーザは、当初グラフィックポインティングデバイスを使用して、ビデオの表示されたフレーム上に矩形領域を作成する。後で定義される領域は先に定義される領域の上に格納され、先に定義される領域と重複することがある。本発明が矩形領域を限定することに関して説明していても、本発明の教示を、円や楕円のような曲線を含む領域、八角形や六角形のような多角形の領域、またはそれ以外の、曲線または直線あるいはその両方を含むユーザが定義する形状にも適用できるのは言うまでもない。ユーザは、各領域を定義してから、領域の優先順位を定義する。代わりに、ユーザは、すべての領域の形状を定義し、その後これら

の領域に優先順位を指定することができる。

【0164】ユーザが初期に領域内で優先順位を定義する場合、領域はピクセル位置に対応する。ただし、以下に示される優先順位を割り当てるプロセスは、マクロブロック単位で動作し、マクロブロックはデジタルビデオエンコードに使用される単位である。したがって、マクロブロックがユーザ定義領域内または外周上にある場合、そのマクロブロックには領域の優先順位が割り当てられる。従来の技術の当業者は、単純な数学を利用してユーザ定義領域に対応するマクロブロックを求めることができる。ユーザ定義領域の外周の場合は、領域の外辺部がマクロブロックに交差する場合に、ユーザ定義領域内に単にマクロブロックを入れる代替策として、マクロブロックを領域内に入れるかどうかを判断するさらに精密なプロセスが、ユーザ定義領域内にマクロブロックの50%以上がある場合には、ユーザ定義領域内にマクロブロックを入れて、ユーザ定義領域内にマクロブロックの50%以下がある場合には、ユーザ定義領域からマクロブロックを排除することによって実行することができる。

【0165】領域およびその優先順位が定義された後で、ユーザは、希望する場合には領域の重複を変更できる。例えば、ユーザは図21の領域469の上で事前に決定されたファンクションキーを押しながら「クリック」して、領域469を領域470の上になるように変更し、人全体に領域469の優先順位が設定され、鳥468だけに領域470に割り当てられる優先順位が設定されるようにする。重複領域には、ユーザによって入力されるときにスタック優先順位が指定される。このスタック優先順位は、ユーザ定義領域が追加、削除、または修正されるときに、必要に応じて調整される。

【0166】領域が定義された後で、ユーザは、希望に応じて領域のけ以上を作りなおしたり、グラフィックポイントングデバイス33を使用して領域を移動することもできるまた、ユーザにより定義された領域は、グラフィックポイントングデバイス33を使用して削除できる。領域が削除されていても、ユーザ定義優先順位だけが削除されているのであり、言うまでもなく削除中の領域内のオリジナルビデオデータは削除されていないことに注記すべきである。

【0167】期間で品質を変更することについての前記項に説明するように、自動エンコードの間または自動エンコードの後で、ユーザはエンコードビデオをレビューすることができ、関心のあるフレームまたは期間が発生するたびに、ユーザはキーボードで単にキーを押したり、入力を別の装置から入力し、特定のフレームまたは期間が重要であることを示す。後で、ユーザはそのフレームまたは期間に戻って、さらに時間を費やして期間を検討し、希望に応じてその期間内で特徴を変えることができる。

【0168】ビデオのフレーム内の領域をユーザが定義

するプロセスは、ユーザによって修正されるデータのフレームが1つしかない単純なケースについて前述された。しかし、ビデオシーケンス内のデータのフレームを一度に1フレームずつ作業するのはきわめて単調で退屈であるため、本発明を使用すると、ユーザは最初の期間で領域を定義し、後の期間で対応する領域を定義できるようになる。それ以降、最初のフレームと最後のフレームの中間のビデオのすべてのフレームに、ユーザにより最初のフレームと最後のフレームに定義される領域に対応する領域が作成され、最初の領域または最後の領域あるいはその両方の領域と同じ優先順位、もしくは最初のフレームと最後のフレームの優先順位に基づく優先順位が指定されるように、補間プロセスが実行される。さらに、重複領域のスタック優先順位は、補間プロセスの間維持される。中間フレームにその領域が自動的に作成された後、ユーザは、希望に応じて、優先順位、領域の形状を変更することにより改変したり、新しい領域を追加したり、領域を削除することができる。

【0169】以下に、補間プロセスがどのように動作するのかの例を示す。ユーザに時間インスタンス t_1 と t_n での2つのフレームの優先順位領域を選択させる。ユーザがフレーム t_1 に領域を、 t_n に対応する領域を割り当てた後、これら2つの領域の座標は、 t_1 と t_n の間のすべてのフレームの領域仕様を獲得するために中間フレーム上で補間される。定義中の実際の領域に補間を実行できるだけでなく、補間は品質優先順位にも実行できる。補間プロセスには、単純な時間関数を使用される。ユーザがポイント a_1 のあるフレーム t_1 の領域Aを定義し、フレーム t_n の対応する領域Aに対応するポイント a_n があり、中間フレーム $a_2, a_3, \dots, a_{(n-1)}$ 内の対応するポイントに、以下のように定義される水平縦軸

$$\begin{aligned} a_i(h) \\ = a_1(h) + (a_n(h) - a_1(h)) \\ \times (i-1) / (n-1) \end{aligned} \quad (23)$$

および以下のような垂直縦軸

$$\begin{aligned} a_i(v) \\ = a_1(v) + (a_n(v) - a_1(v)) \\ \times (i-1) / (n-1) \end{aligned} \quad (24)$$

があると仮定する。この場合 h と v は問題のポイントの水平縦軸と垂直縦軸に対応する。このプロセスでは、領域を限定する多角形の頂点だけが使用され、頂点は補間される。それから、領域は頂点により限定される。線形補間以外の他の補間技法も実行できる。

【0170】前記補間プロセスは、編集対象のセグメント内のすべてのフレームのすべての領域が適切に定義されるまで、必要な回数実行できる。さらにこのプロセスは、ユーザが中間プロセスの補間で決定された領域を移動したり、領域のサイズを変更することができ、補間プ

ロセスが最初のフレームと改変された中間プロセスに、それから中間とフレームと最後のプロセスにもう一度実行できるように発生する微調整プロセスにより再帰的となる。補間プロセスは、おもに、フレームを横切って移動する領域を定義するために使用される。しかし、本発明には、静止しておりフレームの列内で使用される領域を一度定義することも含まれる。

【0171】領域が定義され、ユーザ定義優先順位が入力された後、各マクロブロックの新しい量子化値は、ユーザによって入力された優先順位およびマクロブロックの古い量子化値に基づいて、図20のステップ456で概算されるか、求められる。これは、ユーザ選択優先順位のそれぞれで使用される元の量子化値の端数を示している図22に説明される関係に従って決定される。例えば、ユーザが優先順位ゼロを入力すると、使用される元の量子化値の結果として生じる端数は1であり、元の量子化値に変更がないことを意味する。ユーザが優先順位-3を選択すると、元の量子化値の端数は1.58となり、優先順位3が設定されるマクロブロックの元の量子化値が、結果として生じる量子化値を求めるために、1.58により乗算されることを意味する。優先順位-3は品質の低下を意味するので、量子化レベルは、品質の低下を達成するために量子化レベルを引き上げる、つまりマクロブロックに使用されるビットはさらに少なくななければならない。反対に、3のような正の優先順位がユーザによって入力されると、元の自動号化からのマクロブロックの量子化値は量子化レベルを引き下げる0.42で乗算される。引き下げられた量子化レベルは、マクロブロック内のイメージを表すにはさらに多くのビットが必要とされ、したがって、画質が向上することを意味する。図22に示される線上のポイントは(-5, 100)、(-4, 1.75)、(-3, 1.58)、(-2, 1.42)、(-1, 1.25)、(0, 1)、(1, 0.72)、(2, 0.58)、(3, 0.42)、(4, 0.25)、および(5, 0.01)である。図22に説明されるマッピングは、q-レベルと優先順位の間この方向の線形関係に基づいている。図22で求められたポイントは、実験を通して求められ、さまざまな端数は希望に応じてユーザ選択優先順位と使用できる。

【0172】優先順位-5と5はそれぞれ最低可能品質および最高可能品質を獲得することを目的とする。最低品質は、非線形量子化スケールが使用されるとき最大の許容量子化レベル112、および線形量子化スケールがMPEG-2エンコードに使用されるとき62の結果である。また、最高品質を達成するために使用される最低量子化レベルは1である。したがって、図22に説明される関係からは、実際の許容qスケールから離れた値が生じることがあるが、これらの計算されたq値は単にその飽和レベルで格納されるだけである。例えば、当初、マクロブロックに量子化スケール20が設定され、ユーザがマクロブロックの優先順位を-5となるように選択すると想定してください。100で乗算される量子化値20が、量子化スケールから離れた新しい量子化レベル2,000を生み出すため、最大量子化値または飽和量子化値が使用される。

【0173】前記のように飽和レベルを使用することに加えて、本発明では、MPEGエンコードのIフレーム、Pフレーム、およびBフレームの予防措置を講じる。本発明は、Iタイプフレームの最高量子化レベルを、非線形量子化ケースの場合96となるように、線形量子化スケール使用時には58となるように制限する。同様に、B型フレームおよびP型フレームの最高量子化レベルは、非線形量子化スケールの場合112、線形量子化スケールの場合62である。線形量子化スケールおよび非線形量子化スケールは、MPEG規格で定義される。Iフレーム型、Pフレーム型、およびBフレーム型のそれぞれに使用される最低量子化レベルは1である。言うまでもなく、図22にブロック図で示されるユーザ定義入力および元の量子化値の端数は、ユーザの要件を満たすために修正できる。

【0174】ステップ456でマクロブロックごとに新規量子化値が計算した後、ステップ458で、ステップ456で求められた量氏化器値の結果生じるフレームごとにビット数を求める。本発明の実施例では、他の領域の品質は低下したが、いくつかの領域の品質が向上した後で、各フレームに同じビット数を獲得しなければならないため、このステップは重要である。新しい量子化値の結果生じるビット数を求めるために、マクロブロックごとの結果として生じるビット数に対する量子化レベルを適切に概算する関数を設定する必要がある。結果として生じるビット数と量子化レベルの関係正を求めるために、MPEG-2エンコードのようなビデオエンコードの経験的な結果の検討が実施され、求められた関係は実際の経験的な結果に基づいている。多くのビデオシーケンスを使用するさまざまな量子化レベルおよびフレームタイプのシミュレーション結果が実施された。これらの結果に基づき、マクロブロックの量子化レベル(q)とそのマクロブロックをエンコードする際に使用されるビット数(b)の間の以下の関係性は、フレームタイプに応じて、以下の通りであると判断された。

【0175】

$$\log(b) = m \log(q) + n \quad (25)$$

この場合、Iタイプフレームの場合 $m=-0.75$ で、BタイプフレームとPタイプフレームの場合、 $m=0.10$ である。さらに、Iフレーム、Pフレーム、およびBフレームの場合それぞれ $n=15$ 、14.5、および13.8である。これらの値は、前記に注記されるように、多数のビデオシーケンスで平均化される。前記等式は、エンコードの動作を予測するための優れた近似基準として働き、ビット数と量子化レベルの間の関係はログ領域内で線形である。言うまでもなく、正確であるなら、量子化レベルと結果として生じるビット数の間のそれ以外の関係も、本明細書に記載されるその他のレート-量子化関数を含む、指定され

た量子化レベルから生じるビット数を求める場合に使用できる。

【0176】前記プロセスは、指定された指定量子化レベルから生じるビット数を求める方法を説明する。しかし、複数のエンコードが使用される場合は、新規エンコードからの情報を使用して、量子化関数に対してデータレートをさらに正確にすることができる。初期自動エンコードプロセスの間に、エンコードされるべきビデオの統計を収集するのに使用される第1ビデオエンコードパスがある。それから、実際のエンコードである第2パスの間に、量子化関数に対するデータレートの別の実際のポイントが、指定されたq-レベルから作成されるビット数bを記述する前記等式に対応するqプロットに対するbにおいてのように獲得される。データが、以下に説明するように、ユーザ指定選択に従ってエンコードされた後、結果として生じるビデオは、それが希望される品質の改善を満たすかどうかに応じて、ユーザによって受け入れられるか、拒絶される。再エンコードするたびに、概算されたビットに対するq-レベル関係の実際データが作成される。新しいq-レベル割当てq'が指定されると、求められたデータレートは、再エンコードによって得られる追加データに基づいて、このポイントに関してさらに正確になる。新規量子化レベル、q'が以前に得られた2つのポイントの間に該当する場合、新規q-レベルで出力されるビット数の概算を得るために、ログ領域内で線形補間を実行することができる。新規量子化レベル「q」が2つの事前に決定されたq-レベルの間に該当しない場合には、モデルは、もっとも近いq-レベルから新しい量子化レベルq'に補外し、エンコード時に作成されるビット数を概算するのに使用できる。データが再エンコードされる時間が長いほど、エンコーダからの出力ビットに対してより優れた概算を得る確率が高くなることを注記する。

【0177】概算された量子化値から生じる各フレームのビット数が図20のステップ458で求められた後、ステップ460で、ビット概算数がフレームの元のビット数に十分近いかどうか判断し、必要ならビット数を訂正する。これは、図23-25Cでブロック図で示されるプロセスに従い実行される。

【0178】ステップ460が実行されると、図23のステップ472が、最初に、ビット概算数と元のエンコードのビット数の差異を求める。これは、以下の等式に従い計算される。

【0179】

$$D = B' - B \quad (26)$$

この場合、B'は、修正されたフレームの概算ビット数、Bは元のエンコードから結果として生じるフレームのビット数、およびDは差異である。元のエンコードから結果として生じるビット数以外のBの値を使用することが

できるのは言うまでもない。例えば、ユーザは、フレームの総ビット数の増加を希望する場合、元のエンコードから結果として生じるビット数より大きい事前に決定された量である数にBを設定できる。同様に、ユーザは、フレームの元のビット数を下回るようにBを設定することもできる。したがって、記録されたフレームは、記憶領域の制約、最小品質制約、およびビデオバッファアンダフローの可能性に応じて、元のエンコードより多いビットまたは少ないビットを消費することがある。また、フレーム内の領域の量子化値が改変されていない場合、その量子化値を改変された領域から結果として生じるビット数だけを分析し、変更されなかった領域を無視することができる。それから、ステップ474でこの差異を分析し、ビットが多すぎると判断されると、フローは、ビット数を低減するためにマクロブロックの量子化レベルを引き上げる目的の図24A-24Cに説明されるプロセスのために、Bに進む。ステップ474で修正されたフレーム内のビットが少なすぎると判断されると、新規フレームのマクロブロックの量子化レベルは、図25A-図25Cにブロック図で示されるプロセスFに従ってさらに多くのビットを作成するために引き下げられる必要がある。それ以外の場合、差異が許容できるスレッシュホールド内にある場合は、マクロブロックの量子化レベルを修正する必要はなく、プロセスは図20のステップ462を実行するために戻る。フレーム内のビット数を再エンコードされたビデオに対し同じに保つ場合、ビットは、ある領域から別の領域にシフトされることが考えることができる。

【0180】ステップ474では、差異Dは、差異をスレッシュホールドTに比較することで分析される。スレッシュホールドは、領域の品質の変更の結果生じるビット数が総ビット数の0.001以内である場合、差異は許容でき、プロセスは停止できることを意味する0.001xBに設定される。言うまでもなく、差異が許容できることを示す停止スレッシュホールドは、別の方法で求められ、希望に応じて引き上げたり、引き下げたり、フレームの元のビット数とユーザ選択優先順位から結果的に生じるビット数の両方に基づくことができる。図24A-図24C、および図25A-図25Cに説明されるプロセスは、図24A-図24Bがユーザ選択品質の結果として生じるビット数を低減するために使用され、図25A-図25Cがユーザ選択品質の結果として生じるビット数を増加させるために使用されるという点を除き、きわめて類似している。図24A-図25Cで実行されるプロセスの一般的な概要は、これらのプロセスの詳細を説明する前にここで記述する。当初、ビデオのフレーム内のマクロブロックは、「無定義」、「負の優先順位」、「正の優先順位」、「変更不可」の4種類の内の1つとしてラベルが付けられる。領域にユーザによる優先順位が指定されていない場合、またはシステムによってユーザが「無定義」ステータスを割り当てることができる場合、その領域は「無定義」タイプであると見なされ

る。ユーザが-5と-1の間の、および-5と-1を含む優先順位を割り当てると、これらの領域内のマクロブロックは「負の優先順位」であると考えられる。領域に1と5の間の、および1と5を含む優先順位が割り当てられると、これらの領域には「正の優先順位」が設定されていると考えられる。最後に、優先順位0で割り当てられるすべての領域、またはシステムによってユーザが他のなんらかの方法で、領域の品質を変更させてはならないことを示すことができる場合には、それらの領域のマクロブロックは「変更不可」タイプと見なされる。4種類のマクロブロックのそれぞれの中でのマクロブロックのそれぞれには、4種類の内のそれぞれに独自のインデックスのセットが備えられるように、インデックス i でラベルが付けられる。ゼロのインデックス i は、最初のマクロブロックに割り当てられ、インデックス $n-1$ は最後のマクロブロックに割り当てられる。この場合、 N はある領域種類内のマクロブロックの総数である。例えば、「無定義」領域のすべてに対し、これらの領域内のマクロブロックのそれぞれに0から $n-1$ のインデックスが付けられる。3つ以上の「無定義」領域がある場合、「無定義」領域に2つの異なるインデックス $i=0$ がないことに注記する。すべての「無定義」領域に対しては、インデックス0が付いたマクロブロックは1つだけ、インデックス1が付いたマクロブロックは1つなどである。

【0181】ユーザが領域の優先順位を決定した後にフレーム内のビットが多すぎる場合、フレームの量子化レベルは、フレームのビット数を低減するために引き上げられなければならない。マクロブロックの元の量子化値が、図10(A)または図10(B)のどちらかにブロック図で示されるフォーマットを持つマクロブロックレーヤのログファイルから削除される。代わりに、エンコーダは再実行し、量子化値がエンコーダの元の実行の後に記憶されていなかった場合には、量子化値を求める。しかし、別の代替策では、エンコードされたビットストリームを復号化し、元のエンコードの量子化値を求める。ビット数を低減するために、手順は、「無定義」タイプのマクロブロックごとの量子化レベルを1量子化レベル、一度に1マクロブロックずつ引き上げることで動作する。これで問題が解決されない場合、「無定義」の各マクロブロックの q レベルは、一度に1マクロブロックずつ再度1増加される。事前に決定された最大量子化レベルが設定され、無マクロブロックの量子化レベルを超えて増加できる。「無定義」タイプのすべてのマクロブロックがこの最大レベルに達すると、「負の優先順位」および「正の優先順位」の領域は、同じように1増加される。まず、すべての負の優先順位マクロブロックが1量子化レベル引き上げられ、これでビット差の問題が解決されない場合、「正の優先順位」マクロブロックのマクロブロックの量子化レベルが1増加される。これで問題が解決しない場合、「負の優先順位」マクロブロックの

量子化レベルは再び1増加され、これで問題が解決しない場合には、「正の優先順位」タイプのマクロブロックの量子化レベルが1増加される。このプロセスは「負の優先順位」タイプおよび「正の優先順位」タイプのすべてのマクロブロックの量子化レベルが前記最大レベルに引き上げられるまで続行する。これでビット差問題が解決しない場合には、「変更不可」タイプのマクロブロックの量子化レベルが、ビット差問題が解決されるまで、一度に1量子化レベル引き上げられる。十中八九、ビット差問題は、「変更不可」タイプの量子化レベルが変更され始める前に訂正されるだろう。ユーザ選択優先順位が設定されるフレームのビット数が少なすぎて、図25A-図25Cのプロセスに従ってビット数を増加するために、マクロブロックの量子化レベルを引き下げ必要がある場合、類似したプロセスが発生する。

【0182】今度は、生じたビットが多すぎてしまったときに図23によって実行される図24Aにブロック図で示されるフローチャートを参照すると、プロセスはブロックタイプを「無定義」タイプに設定するステップ480により図24Aで開始する。それから、ステップ482で、マクロブロックインデックスを、インデックス i がゼロに等しく設定されることを意味する開始ポイントに設定する。それから、設定されていたマクロブロックタイプ（フローチャートで初めて、「無定義」タイプの第1マクロブロック）の問題のマクロブロック（インデックス $= i$ 内のマクロブロック）の量子化レベルに最大事前定義量子化レベルが設定されているかどうかを判断する、ステップ486が実行される。この訂正プロセスの最大量子化レベルは、訂正されるマクロブロックの品質で大きな劣化がないように高すぎて設定してはならない。可能であるなら、マクロブロックの質に大きな差が出ないように、均一のビット数を減少させる目的で品質を劣化させようとするのが望ましい。したがって、この最大値は線形量子化スケールが使用されるときには量子化レベル62に、非線形量子化スケールがMPPEG-2エンコードなどに使用されるときには112に設定される。ただし、最大量子化器値には、それ以外の値も使用できる。問題のマクロブロックの量子化レベルがステップ486で最大ではないと判断されると、ステップ488で問題のマクロブロックの量子化レベルを1量子化レベル増分する。それから、ステップ490で前記のビットレート-量子化機能を使用してフレームのビット概算を計算し直す。それからステップ490では、マクロブロックのインデックスをゾウ文する。マクロブロックのインデックスがマクロブロックタイプの最後のマクロブロックが処理されていないことを示す場合は、フローはステップ484にループバックする。それ以外の場合は、フローはステップ492から、ブロックインデックスをゼロにリセットするステップ498に進み、第1マクロブロックを示す。

【0183】フローがステップ484に戻ると、ステップ4

90で計算し直されるビット概算は、前述するように、スレッシュホールドに比較される前記の差Dを計算し直すために使用される。ビット概算が高すぎない場合、ユーザ選択優先順位を指定されるフレームのビット数が訂正され、フローは呼び出しプロセスに戻る。それ以外の場合、フローは、問題の（インデックス1が設定される）マクロブロックの量子化値が最大であるかどうかをもう一度判断するステップ486に進む。

【0184】ステップ486で、問題のマクロブロックの量子化値が最大であると判断すると、そのブロックの量子化レベルは引き上げる必要はなく、ステップ494では1ブロックインデックス*i*を増分する。ステップ496でマクロブロックタイプのインデックスが最後のブロックを通り過ぎていないと判断すると、フローはステップ486に進み、増分されたインデックスの付いたブロックに最大量子化値が設定されているかどうかを判断する。それ以外の場合、ステップ496でマクロブロックが問題のマクロブロックタイプの最後のマクロブロックであると判断すると、ステップ498が実行され、ブロックインデックスが第1マクロブロック（*i*=0）にリセットされる。それから、フローは図24Bに説明されるプロセスCに進む。

【0185】図24Bでは、ステップ500で、ブロックタイプが「無定義」タイプであるかどうか調べる。「無定義」タイプである場合は、ステップ502ですべての「無定義」マクロブロックに最大qレベルが設定されているかどうかを判断する。「無定義」マクロブロックのすべてに最大qレベルが設定される場合、「無定義」タイプのマクロブロックにさらに調整を実行することは不可能であるため、マクロブロックタイプは、ステップ504で「負の優先順位」に変更され、フローは図24Aにブロック図で示されるプロセスDに戻り、ステップ484が実行される。それ以外の場合、ステップ502で、「無定義」タイプのすべてのマクロブロックに最大qレベルが設定されていないと判断すると、フローは修正中のマクロブロックのタイプを変更しないで、プロセスに戻る。この手順は、フレームの結果として生じる数がスレッシュホールド範囲内になるまで続行するか、あるいは「無定義」型のすべてのマクロブロックが「負の優先順位」に変更されるブロックタイプを結果的に生じさせる最大量子化レベルに設定される。

【0186】ステップ500でブロックタイプが「無定義」ではないと判断するとステップ506がブロックタイプが「負の優先順位」であるか判断する。ブロックタイプが「負の優先順位」の場合、ステップ506は、「正の優先順位」のすべてのブロックに最大qレベルが設定されるかどうか判断する。ステップ508の判断の結果が負である場合、ステップ510が実行され、ブロックタイプが「正の優先順位」に設定され、フローは図24Aにブロック図で示されるプロセスDに進む。

【0187】ステップ508で「正の優先順位」のすべてのマクロブロックに最大qレベルが設定されると判断すると、ステップ512で「負の優先順位」タイプのすべてのブロックに最大qレベルが設定されるかどうか調べる。設定されていない場合、ブロックタイプは変更されず、フローは図24AのプロセスDに戻る。「負の優先順位」タイプのすべてのマクロブロックに、ステップ512で、最大qレベルが設定されていると判断される場合、すべての「無定義」、「負の優先順位」および「正の優先順位」のマクロブロックに最大量子化レベルが設定され、ブロックタイプがステップ512で「変更不可」に設定される。それから、フローは図24AのプロセスDに進む。

【0188】図24Bのステップ506が負と判断されてから、プロセスEおよび図24Cのステップ520が実行される。ステップ520でブロックタイプが「正の優先順位」であるかどうか判断する。そうである場合は、ステップ522で「負の優先順位」のすべてのブロックに最大qレベルが設定されるかどうか判断される。設定されていない場合、ブロックタイプは「負の優先順位」に設定され、フローは図24Aにブロック図で示されるプロセスDに戻る。ステップ522で「負の優先順位」タイプのすべてのブロックに最大qレベルが設定されていると判断すると、ステップ526で「正の優先順位」タイプのすべてのブロックに最大qレベルが設定されるかどうかを調べる。設定されていない場合、フローは図24Aにブロック図で示されるプロセスDに戻る。それ以外の場合、すべての「無定義」ブロックとともにすべての「負の優先順位」および「正の優先順位」に最大量子化レベルが設定され、ステップ526でブロックタイプが「変更不可」に設定され、フローは図24Aにブロック図で示されるプロセスDに戻る。

【0189】ステップ520でブロックタイプが「正の優先順位」ではないと判断すると、ブロックタイプはそれゆえ「変更不可」でなければならず、ステップ530で「変更不可」タイプのすべてのブロックに最大qレベルが設定されるかどうか判断される。設定されている場合、ユーザ選択優先順位が設定されるフレーム内のビットが多すぎるという問題を解決しないで、すべてのブロックタイプが最大量子化レベルに設定されたので、エラーが生じる。すべてのブロックを最大事前定義qレベルに設定すると、フレームの元のエンコードを上回らないビット数が生じるはずである。ステップ530での判断が負である場合、フローは図24Aに説明されるプロセスDに戻る。

【0190】図23のプロセスが、ユーザ選択優先順位領域のあるフレーム内のビット数が少なすぎると判断すると、図25A-図25Cのプロセスが実行される。図25A-図25Cは、量子化レベルがフレームのビット数を増加させるために引き上げられる代わりに引き下げられ、マクロブロックの量子化レベルが、引き下げられたときに量子化R

レベルの1のような事前に決定される最小レベルを超えることができないという点を除き、図24A-図24Cにブロック図で示されるプロセスに基づいている。前記に注記した相違点を除き、図25A-図25Cのプロセスが図24A-図24Cのプロセスに同一であるため、図25A-図25Cのさらなる説明は簡略を期すために行わない。

【0191】図23-図25Cに記述される訂正プロセスは、本発明が動作する1つの様式である。ただし、異なった種類のマクロブロックの量子化レベルを調整する代替手段も可能である。例えば、「負の優先順位」および「正の優先順位」のマクロブロックの量子化レベルを、前述のように同時に変更する代わりに、「負の優先順位」タイプ、「正の優先順位」タイプおよび「変更不可」タイプのマクロブロックは均一に増加できる。代わりに、ビットが多すぎる場合には、「負の優先順位」タイプの量子化レベルが、最大量子レベルに達するまで引き上げられてから、「正の優先順位」タイプのマクロブロックの量子化レベルが引き上げられる。反対に、ユーザ選択優先順位が設定されるビット数が少なすぎる場合、「負の優先順位」タイプのマクロブロックのレベルが引き下げられる前にビット数を増加するために、「正の優先順位」タイプのマクロブロックの量子化レベルが引き下げられる場合がある。後者の2つのケースでは、「無定義」タイプのマクロブロックの改変は、「負の優先順位」または「正の優先順位」タイプのマクロブロックの前、間、または後に発生する可能性がある。

【0192】図23A-25Cのプロセスが実行された後、フローは、新しい量子化値を使用してビデオを再エンコードする図20のステップ462に戻る。この再エンコードは、再エンコードされたフレームの結果として生じるビット数を、フレームの元のエンコードのビット数にできる限り近づける目的で、レート制御関数を使用して、量子化値を微調整する間に実行される。レート制御装置プロセスは、以下の別項に詳しく説明する。また、再エンコードの間、エンコード人為構造が発生する可能性があるため、以下の別項に説明するように、これらのエンコード人為構造を低減する目的で特殊な手段を実行できる。

【0193】空間手動編集のプロセスは、本発明の実施例の図1Aのワークステーション30で発生する。しかし、空間手動編集がビデオエンコードプロセスと同時に発生しない場合、ワークステーション10は、ビデオエンコードプロセスのレートを劣化させないで時間的手動編集を事項できる。本発明がフレーム内のさまざまな領域の画質を変更することとして詳しく記述されることに注記すべきである。本明細書の概念を、あるフレームの領域内のビットを別のフレームの領域にシフトすることに拡大することができるのは言うまでもない。さらに、空間手動編集は、前記の時間的手動編集といっしょに適用できる。

【0194】空間手動編集は、ユーザ指定領域優先順位が設定されるフレームのビット数をフレームの元のエンコードと同じに保つこととして説明された。ただし、希望される場合、再エンコードされた編集セグメントは、十分な記憶領域容量がある場合には、フレームの元のビット数より多いか少ない事前に決定されるビット数を消費できる。これらのケースでは、改変されたフレームを変更できるビット数を理解しておくことが望ましく、そのため図24A-図25Cにブロック図で示されるプロセスは、設定された差が達成されるまで動作するだろう。したがって、任意のフレームの再エンコードのビット総数は、5%と1%多いまたは少ないビットの両方を含む0%と20%の間で多くまたは少なく、元のエンコードを上回るか、下回る任意の量を消費する。

【0195】空間手動編集の説明は、ビデオがエンコードされた後のビデオの変更に際して行われた。しかし、本明細書に記載される教示は、初期エンコードプロセスがなかったシステムにも適用できる。

【0196】空間手動編集は、以前に取得されたフレームと同じビット数を達成するために、ユーザ選択優先順位に基づいて量子化値を割り当ててから、必要に応じて、フレームのマクロブロックの量子化レベルを調整することによって動作する。対照的に、前記の項に記述された時間的手動編集は、ビデオのセクションからビットの共通プールを削除し、それらのビットを、ユーザ選択優先順位に従って、共通プールから分配し直す。あるフレームに関して空間手動編集が実行されるプロセスについてのすべての教示は、時間的手動編集の項に説明されるように、時間でビデオのフレームの品質を変更するという概念に適用でき、逆の場合も同じである。例えば、空間編集は、品質の変更を達成するために量子化レベルを修正することとして説明されてきたが、ビット数は時間的編集の項で調整される。フレームまたはマクロブロックのビット数は、直接、量子化レベルに依存するため、ビット数および量子化レベルの修正は、同じ結果を達成し、そのため時間的編集および空間編集は、ビット数または量子化レベルあるいはその両方を修正できる。

VIII. 品質変更後のビデオの再エンコード

A. 品質が単独フレーム内で変更された場合のビット割当て制御

図1Aにブロック図に示されるビデオエンコード装置50のビットレートを制御するために、ワークステーション10または同じ制御機能を実行できるあらゆる専門ハードウェアのような制御装置が、量子化レベル、つまりq-レベルおよびレート制御式エンコーダ内へのスライスごとの概算ビット数または予想ビット数をロードする。それから、エンコーダはエンコードを開始し、量子化レベルは、以前に求められた予想ビット数に比較して、作成されたビット数の精度を測定してから調整される。予想ビット数は、空間手動編集および時間的手動編集に関する

項で前述されるように、レート量子化関数を使用して求められる。このプロセスは、一度の n 個のマクロブロックの集合に対して連続して実行され、各集合は $MB(i)$ と示され、以下の通りである。

【0197】

$1 \leq i < (\text{マクロブロックの合計数}/n)$ (27)

マクロブロックの最後の集合を再エンコードする場合、作成されるビット数のマイナーな過剰予想が、ビットスタッフィングを使用して吸収できるため、レート制御装置が、再エンコードプロセス中で作成されたビットの総数が最初にエンコードされた長さより短くなることを保証する。

【0198】図26は、事前に割り当てられる量子化レベルでデジタルビデオの集合を再エンコードするための方法を説明するフローチャートである。調整は、量子化モデルが、再エンコードにより作成されるビット数を、作成されたビットの実際の数に対してどれほど正確に予想下のかに基づいて、事前に割り当てられた量子化レベルに加えられる。図26では、レート制御式コーディングプロセスは、ステップ700で開始してから、ステップ702で $i=1$ を設定することによって、 n マクロブロックの第1集合をマクロブロックの現在の集合として指定する。ステップ704では、マクロブロックの各集合 $MB(i)$ の再エンコードで使用される、ビット予想数の完全な集合、 $EB(i)$ をロードする。それから、 $MB(i)$ はステップ706で再エンコードされ、実際に作成されたビットの数が $GB(i)$ として記憶される。作成済みビット $GB(i)$ 、予想ビット $EB(i)$ の間のパーセンテージ差は、ステップ708で予想率、 $PB(i) = GB(i) / EB(i)$ として計算される。さらに、システムは、累積ビットの予想数 $EB'(i)$ 、累積作成ビット数、 $GB'(i)$ をマクロブロックのすべてお再エンコードされた集合に対して計算することが可能で、この場合、 $EB'(1) = EB(1)$ 、 $GB'(1) = GB(1)$ 、 $EB'(i) = EB'(i-1) + EB(i)$ および $i > 2$ の場合、 $GB'(i) = GB'(i-1) + GB(i)$ である。したがって、 $PB(i) / EB'(i)$ を使用することによって、説明されたステップ708の代替策として、 $PB(i)$ は、累積ビット予想数 $EB'(i)$ および累積作成ビット数 $GB'(i)$ を使用して計算できる。これは、マクロブロックのある特定の集合がどれほど正確にエンコードされたのかを示す代わりに、全体としての再エンコードプロセスがどのようにターゲットビット数を満たしているのかのより正確な表示となる。ステップ710では、概算率を使用され、マクロブロックの次の集合、 $MB(i+1)$ 内の q -レベルの訂正係数を判断する。

【0199】この訂正係数は、2通りの方法の内の1つを使用して求められる。第1の訂正係数決定方法は、割合ルックアップテーブルと呼ばれる割合値のテーブル内の率の単純並べ替え済みリストを進み、どの2つの割合オ間に概算率が該当するのかを判断する。テーブル内の2

つの数の大きい方のインデックス j は、訂正係数インデックスとして使用される。訂正係数インデックスは、テーブル内の j 番目の要素を訂正係数として選択することにより訂正係数テーブル内で訂正係数 $\Delta(i)$ を探すために使用され、マクロブロックの次の集合は $\Delta(i)$ によって調整される。割合ルックアップテーブル、 RT には、例えば、以下のような並べ替えられた値が指定され
 $RT = \{0.85, 0.90, 0.95, 1.0, 1.05, 1.10, 1.15, 1.2, 1.5\}$

訂正係数テーブル CT には、以下の値が指定される。

【0200】

$CT = \{-3, -2, -1, 0, 1, 2, 3, 4, 9\}$

図30は、概算率を使用して訂正係数インデックスを割合ルックアップテーブルから求めるプロセスを説明するフローチャートである。ブロック図のためだけに、 $PB(i) = 0.87$ であると想定する。インデックス j は、当初、ステップ722で設定される。 $PB(i)$ を $RT(j)$ に比較するプロセスが開始し、前記ルックアップテーブルに定義されるように、ステップ724では、 $PB(i) = 0.87$ が $RT(1) = 0.85$ に比較される。0.87が0.85を上回る場合、プロセスはステップ726に続く。ステップ726で残りのインデックスがあると判断されているので、制御はステップ728に渡され、 j は1増分される。ステップ724に戻ると、 $PB(i) = 0.87$ は現在 $RT(2) = 0.90$ であるため、制御は、訂正係数インデックスの検索を終了するステップ730に渡される。

【0201】図31は、テーブルへのインデックスが指定されるルックアップテーブル内で値を捜し出す方法を説明するフローチャートである。図31では、第1の方法に従って訂正係数を求めるために、2であるインデックス j が使用され、 CT 内の対応するエントリを捜し出す。 CT 内の第2位置から読み取ると、訂正係数は、ステップ736で-2であると判断される。したがって、 $MB(i+1)$ の q -レベルは-2で変更される。同様に、 $PB(i) = 1.12$ の場合、インデックス j は $PB(i)$ より大きい RT 内の最初の入力に対応するため、7番目の位置にある CT の訂正係数は3である。したがって、 $MB(i+1)$ の q -レベルは3を追加することによって、例えば3を $MB(i+1)$ に加算することにより変更される。

【0202】第2の訂正係数決定方法では、再エンコードされたマクロブロックの集合のパーセンテージが訂正係数を求める際に第2パラメータとして使用される点を除き、第1方法に類似したプロセスを使用する。さらに、テーブルから訂正係数を読み取る代わりに、訂正係数はマトリックスから読み取られる。再エンコードされたマクロブロックの集合のパーセンテージ、または再エンコードされていないマクロブロックの集合のパーセンテージを使用することにより、フレームの列の最初でより漸進的な訂正を加え、必要なら、最後でより著しい訂正を加えることができる。これによって、システムは量

子化モデルと実際に作成されたビット数の間の変動をより正確に訂正できるようになる。再エンコードされずに残っているマクロブロックの集合の数と訂正の意義の間の反転関係を使用し、マクロ部アロケーションのある集合内での過剰予想がマクロブロックの別の集合での過剰予想により相殺されると仮定すれば、不必要な訂正は回避され

$$CTM = \begin{bmatrix} -1 & -1 & 0 & 0 & 0 & 1 & 1 & 2 & 4 \\ -2 & -1 & 0 & 0 & 0 & 1 & 1 & 2 & 5 \\ & & & & & & & & \\ & & & & & & & & \\ -3 & -2 & -1 & 0 & 1 & 2 & 3 & 4 & 8 \\ -3 & -2 & -1 & 0 & 1 & 2 & 3 & 4 & 9 \end{bmatrix}$$

【0204】フレームの列の最初で、概算比が必要なビット数の過剰評価を示す場合、すべての未使用のビットをただちに次のスライスに割り当てし直す必要はない。未使用のビットがただちにMB(i+1)によって消費された場合、k i+1の場合、予想以上のビットを消費するMB(k)により、MB(k+1)は強制的に不必要にさらに高い量子化レベルに引き上げられるだろう。未使用のビットは、必要とされ、MB(k)によって使用されるまで、再エンコードプロセスでMB(i)から前方に運ばれる。

【0205】図32は、概算比および残っている再エンコードされていないマクロブロックの集合の数の両方を使用して訂正係数がどのように計算されるのかを説明するフローチャートである。必要となるビット数の過剰評価のケース、つまり訂正係数jがPB(i)=0.87の場合に2に等しいと以前に判断した場合、図32のステップ744で過剰評価が再エンコードプロセスの最初のパーセントデ発生するかどうか判断する。CTMの(2,1)での要素は-1であり、したがって訂正係数として選択される。これにより、第1方法では-2がであった再エンコードプロセスの初期に訂正効果は鈍る。しかし、再エンコードプロセスの99パーセントは実行され、概算比がPB(i)=1.12である場合、3である要素(9,77)が選択されるので、再エンコードプロセスの最後で完全に訂正する。

【0206】図27(A)-図27(C)に戻ると、q-レベルが事前に割り当てられたフレームがある集合内のマクロブロックの数が1つのスライスの中のマクロブロックの数に等しい別の例として説明される。ビットの概算数EB

(i)が前述のように計算され、図27(A)のMB

(i)、マクロブロックiのレート制御プロセスで使用されるメモリの中にロードされてから、MB(i)が実際に再エンコードされる。作成ビット数、GB(i)が求められ、概算比が計算され、訂正係数探し出される。図27(B)に示されるように、それからMB(i+1)は△

(i)によって訂正され、プロセスはMB(i+1)のために繰り返される。図27(C)には、MB(i+1)が概算され、計算され、その概算比が求められた後のMB(i+2)の変化が示される。このプロセスは、すべてのマクロブロックおよび部分的なマクロブロックが再エンコードさ

る。

【0203】第2訂正係数方法の訂正係数マトリックスCTMの例として、CTMが以下のように定義されると想定する。

【数1】

れ、訂正されるまで続行される。

【0207】図27(A)および図27(D)を参照して別の例が示される。この例では、集合あたりのマクロブロック数は、2つのスライスの仲のマクロブロックの数である。EB(i)、GB(i)、およびPB(i)を計算した後で、2つのスライスを構成するMB(i+1)のマクロブロックは、図27(D)に説明されるように、△(i)により訂正される。この再エンコードおよび訂正プロセスは、すべてのマクロブロックが再エンコードされるまで以前の例で続行される。

【0208】再エンコードプロセスにより生じる変更は、図28(A)および図28(B)に説明される。フレームNの2つの等しい面積の領域である領域Aと領域Bは、当初ほぼ等しい数のビットを使用してエンコードされる。ユーザは、領域Aの品質が劣化され、領域Bの品質が向上されなければならないと指定する。しかし、元のエンコードに使用されたビット数が再エンコードのビット数にできる限り近くなるようにすることが望ましい。フレームNを再エンコードした後、フレームNには元のエンコードから作成された同じバイト数が格納されるが、領域Aはより少ない数のビットでコーディングされ、領域Bは追加ビットを使用する。これにより、ユーザが要求したように、領域Aの品質は劣化し、領域Bの品質は向上する。本発明のレート制御プロセスでは、レート量子化関数を使用して求められた概算数でのマイナーな誤差を訂正することにより、フレームの再エンコードされたビット数はできる限り元のビット数に近づけられる。

【0209】図29では、複数のフレームからのビットを割り当てし直し、それらを第2の複数のフレームに与えることにより、デジタルビデオストリームに加えられる変更を説明する。個々のフレーム面積は変化しても、フレームの集合の総面積は同じままでなければならない。レート制御装置により、複数のフレームの再エンコードのビット数は、ビット概算数を求めるために使用されるレート量子化関数がわずかに間違っているても、元のエンコードからのビット数と同じになるように微調整できるようになる。

B. 編集ポイントで復号化人為構造を回避しながら再エ

ンコードする

1. 最初にエンコードしたときに使用された状態にエンコードを復元する

前述の時間品質編集動作および空間品質編集動作は、ビデオの品質を新規量子化値でビデオを再エンコードすることによって変更できるようにする。ただし、再エンコードされたビデオを単にオリジナルビデオに代入するだけでは、ビデオの編集ポイントでの可視グリッチとして出現する受け入れられない復号化人為構造が生じる場合がある。これらの可視グリッチは、代入されている新規セグメントにMPEG-2構文のような適切な構文が設定され、編集境界でのバッファ制約事項が満たされていても発生する。問題は、編集セグメントの前に来る最後の基準フレーム（最後のP-フレーム）および編集セグメントの最後にある別の基準フレーム（私語のP-フレーム）の非可用性の結果生じる。これにより、編集セグメントの最初にあるB-フレームの第1集合と編集セグメント直後B-フレームの第1集合のエンコードと復号化の間に格差が生じる。

【0210】デジタルエンコード圧縮ビデオの編集を適切に実行するためには、まず、希望の編集ポイントに対応するエンコードビットストリームの適切な位置を決定する必要がある。このプロセスは、エンコードされたビデオにアクセスするためのユーティリティを説明する第VI項に前述された。前述するように、これは、問題のピクチャのビットオフセットを求めるために、既にエンコードされたフレームのそれぞれい使用されるビット数を合計することで実行される。代わりに、エンコードされたビットストリーム内のある特定のフレームや期間の厳密なロケーションを示す情報を維持するために、ディレクトリも使用できる。

【0211】今度は図33 (A) -図33 (E) を見ると、本発明の動作を説明するための例として使用されるビデオの期間がブロック図されている。図33 (A) では、エンコードされたビデオの元のセクション0がブロック図で示される。エンコードされた期間5 m 0sから10m 0s (5分、ゼロ秒から10分、ゼロ秒まで) を別のビデオセクションで置き換えることが希望される。図33 (B) では、図33 (A) のエンコードされたビデオに代入しなければならない未エンコードビデオの新しいセクションがブロック図で示される。図33 (B) および図33 (C) の破線は、未エンコードビデオを表すのに使用され、図33 (A) 、図33 (D) および図33 (E) の実線はエンコードされたビデオを示すのに使用される。

【0212】図33 (B) に示される新しいビデオのセクションだけが、GOPが以前のGOPを参照する標準MPEG-2を使用してエンコードされ、図33 (A) に示されるビデオに代入される場合、復号化人為構造が、MPEG-2ビデオ圧縮で使用される双方向復号化に必要な正確な基準フレームが欠如しているため、編集ポイント5m 0sと10m 0sで

存在するだろう。GOPの始まりは以前のGOPに戻って参照するが、GOPの最後は次のGOPに進んで参照しないという事実のため、この問題は本発明により違うように処理される5分ポイントおよび10分ポイントの両方で存在する。

【0213】復号化人為構造をGOPの始まりで防ぐには、GOPの最初の2つのB-フレームの基準フレームとして使用される先行するGOPの最後のP-フレームを作成し直すことである。これは、例えば、オリジナルビデオの1つのGOPを挿入対象のビデオの始めに格納することにより達成され、この追加されたビデオは図33 (C) の4 m 58.5sから4 m 59sに示される。例では、GOPの他の期間も可能ではあるが、あるGOPは1秒の2分の1に等しい。エンコードは、4m 59sの直前にGOPの最後のP-フレームを作成するために、元のエンコードに使用される同じ量子化値でこのセクションに関して再実行される。最後のP-フレームを除く、ビデオのこの短いセクションの再エンコードの結果生じるすべての情報は廃棄されることがある。

【0214】編集セグメントの最後での復号化人為構造を防ぐために、オリジナルビデオの1秒という期間（例えば、図33 (C) の10 m 0sから10m 1s）が置換される新規ビデオの最後に付加され、この1秒セグメントは、元のエンコードの間に使用されたのと同じ量子化値を使用してエンコードされる。この1秒期間がビデオの最後に付加されなかった場合、10m 0sマーク直後の2つのB-ピクチャは、10m 0sは2つのGOPの間の分割ポイントであると仮定し、編集の前と編集の後で異なる基準ピクチャに戻って参照する。例は、ビデオの2分の1秒期間または1秒期間の新規セクションへの付加として説明したが、概念はGOPの付加に一般化できる。

【0215】現在では、最後にある編集ポイントが10m 1sポイントであるため、復号化人為構造は10分マークでは出現なくなるが、人は復号化人為構造が10m 1sで生じるかどうかを考慮しなければならない。10m 1sマーク直後の2つのB-フレームが、それらが以前に参照した本質的に同じ基準フレームに戻って参照するため、可視復号化人為構造は生じない。10m 0sマーク直後の2つのB-フレームが、それらが過去に参照した同じ基準フレームに戻って参照しなくても、2つのB-フレームは再エンコードされていたため、適切に新しい基準フレームに戻って参照する。したがって、可視復号化人為構造は10m 0sポイントまたは10m 1sポイントでは存在しない。

【0216】図33の最後に追加された2つのGOP (1秒) は、最後のP-フレームが元のエンコードにほぼ同等となるのに十分な時間でなければならない。10分ポイントの後のオリジナルビデオを再エンコードするときにまったく同じ量子化値が使用されることが保証できる場合には、1GOP (2分の1秒) で正確にエンコードされた基準ピクチャを適切に作成するのに十分であるはずである。た

だし、適正な品質を保証するためには、2つのGOPが好まし。最初に使用されたのと同じ量子化値を使用することが希望されても、実際には、元の量子化値からマイナーな偏差が生じうる場合があるため、2つのGOPが望ましい。

【0217】4m 58.5sから4m 59sおよび10m 0sから10m 1sまで使用されるビデオの元の期間に加えて、図33

(C)は、4m 59sから5m 0sまでの追加オリジナルビデオを示している。このビデオは復号化人為構造の防止には必須ではないため、最後のP-フレームを決定するための2分の1秒が5m 0sマークの直前(例えば4m 59.5sから5m 0sまで)に使用できる。ただし、4m 59sから5m 0s期間は、オペレータによる、品質の手動変更が編集境界を超えて影響したという認知の機会を排除するために新規ビデオに付加される。この誤認は、MPEG GOPの最初でのフレームのコーディング順序はIBBだが、フレームはBBIとして表示されるという事実のためである。1秒期間は、その期間の元のエンコードの間に使用されたのと同じ量子化値を使用してエンコードされる。以前使用された量子化値を得る方法を以下に説明する。1秒期間が4m 59sから5m 0sに付加されないと、図33(D)にブロック図で示される代わりのエンコードされたビデオは、5m 0sで開始するだろう。

【0218】図33(C)のビデオをエンコードした結果生じる有効なデータは、4m 59sから10m 1sまで実行する代替データSとして図33(D)に示される。それから、この代替データSは、図33(E)に示すような最終ビットストリームを結果的に生じさせる図33(A)にブロック図で示される元のエンコードされたデータに代入される。

【0219】図33(A)に示されるエンコードされていない情報およびエンコードされた情報は、実施例のビデオであるが、エンコードされた情報を復号化するために、以前のまたは以降の、あるいはその両方の情報を使用するオーディオまたはその他の種類の情報となる場合がある。

【0220】図33(B)のビデオの新しいセクションは、任意のビデオソースから得られ、前記の時間的編集プロセスまたは空間編集プロセスによって作成されるビデオである可能性がある。4m 59sから5m 0sまでおよび10m 0sから10m 1sまでの図33(C)のオリジナルビデオは、元のエンコードに使用されたのと同じ量子化値が設定されるオリジナルビデオである。図33(C)の新しいビデオNの両端にあるこのビデオは、時間的編集プロセスを使用し、最後のセクションの品質を、時間的手動編集の項に説明した保護つきに設定するだけで得られ、量子化値が同じで留まる必要があることを示す。代わりに、図33(A)のビデオは、図33(A)のビデオに関係なく完了され、異なった長さとなる場合がある。さらに、ビデオのセクションをカットアウトし、新しいビデオを付加しないために本発明を使用することもできる。この

場合、図33(B)にビデオは存在せず、「N」セグメントとその対応するエンコードは、図33(C)-図33(E)には記載されていない。編集セグメントの前の最後のP-フレームが適切に構築され、オリジナルビデオの2分の1秒から1秒の(1つまたは2つのGOP)がエンコードされるべき新しいビデオの最後に付加され、元のエンコードで使用された同じ量子化スケールでエンコードされる限り、可視復号化人為構造はどのような状況でも発生しない。

【0221】図34は、図33(A)-図33(E)に関して記述されるプロセスを説明するフローチャートである。開始後、ステップ780で、ビデオは圧縮フォーマットにエンコードされ、図33(A)に示されるエンコードされたビデオが作成される。ステップ782では、エンコードされたビデオに代入されなければならないエンコードされていないフォーマットのビデオの新しいセクションが決定される。ビデオのこの新しいセクションは、図33

(B)と図33(C)でNとして示され、オリジナルフレームに関連する、オリジナルフレームに関連しない、またはオリジナルフレームとは異なった期間がセ設定される場面を表すことがある。置換されるビデオの開始ポイントおよび終了ポイントは、それぞれ5m 0sと10m 0sである。ただし、置換対象の最終エンコードの実際の開始ポイントおよび最終ポイントは、図33(D)にブロック図で示されるように異なっている場合がある。

【0222】ステップ784では、編集ポイント5m 0sおよび10m 0sでの復号化人為構造を防止するために、前記の理由からエンコードされていないビデオの期間をビデオのエンコードされた新しいセクションの最後に付加する。これらの付加された期間は4m 59sから5m 0sおよび10m 0sから10m 1sまでである。ただし、4m 59sから5m 0sまでの期間は必須ではない。

【0223】ステップ786では、編集ポイントの直前のP-フレームを決定する。図33(E)で置換されるビデオは4m 50sから10m 1sなので、4m 50sの直前の最後のP-フレームを突き止める必要がある。このP-フレームは、4m 58.5sから4m 50sまでの期間のためのエンコードを、P-フレームを最初にエンコードするために使用された量子化値で実行することにより決定される。量子化値は、図10(A)または図10(B)のどちらかに示されるフォーマットを持つマクロブロックログファイルから元の量子化値を検索することにより求められる。ただし、映画全体の各マクロブロックの量子化値は、大量の記憶領域容量を消費するので、量子化値を記憶するのは望ましくない場合がある。代替策として、量子化値を、Iフレーム、Pフレーム、およびBフレーム用ビデオバッファ、allocated_bitとして図9Bにピクチャレイヤログファイルに図示されるピクチャのターゲットビット、図9Cにmean_activityとして記憶されるピクチャの活動レベル、およびエンコードの間にエンコードにより計算されるマクロブロックの活動レベルに関するレート制御ステータスを使用し

て単に作り出すことができる。レート制御ステータスは、それぞれIビデオバッファ、Pビデオバッファ、およびBビデオバッファのレート制御ステータス情報を記憶する、図9Cの底部でS1_i、S2_i、S1_p、S2_p、S1_b、およびS2_bとしてブロック図で示されるピクチャレイヤログファイルに記憶される。ビデオエンコード技術の当業者は、本明細書に記載される教示に基づき、過度の実験を行わないで希望のP-フレームの量子化値を求めることができるだろう。

【0224】基準フレームの量子化値またはエンコードされた基準フレーム自体を求める代わりの方法として、エンコードされた基準ストリーム内の基準フレームのエンコードバージョンは既知の方法で複合化できる。さらに、必要とされる基準フレームを得るそれ以外の方法も実行できる。

【0225】基準P-フレームが得られた後、ビデオは、最後のセクションの元のエンコードと同じ量子化レベルを維持しながら、ステップ788でエンコードされる。量子化レベルは、前述のように得られる。結果として生じるエンコードされたビデオは、図33(D)に示される。

【0226】それから、ステップ790で、新規にエンコードされたビデオを最初にエンコードされたビデオに代入し、図33(E)に説明するように、最終製品が得られる。置換されたエンコードされたビデオを含むこのエンコードされたビデオには、編集ポイントで最小の復号化人為構造がある。

2. ピクチャのグループの閉鎖モードでの再エンコード
図33(A) - 図34に説明される手順の代替策として、復号化人為構造を低減しつつも新しいビデオの始まりのエンコードのために別の手順を実行できる。この手順では、ビデオの立ち下がり端を処理する方法は、図33(A) - 図34に関して前記と同じである。

【0227】この代替実施例の例として、図33(A)を元のエンコードされたビデオとして、図33(B)を図33(A)に代入されなければならないエンコードされていないビデオとして使用する。元のエンコードされていないビデオは図33(B)の始まり縁で追加されず、図33(C)に記述されるように、対応する元のエンコードされていないビデオの2つのGOP(1秒)は、図33(B)の最後に追加され、エンコードされるべきビデオは図35(A)に示される通りである。

【0228】エンコーダが、第2基準ピクチャのないB-フレームのビットを増加させながら、図35(A)のビデオをGOP閉鎖モードでビデオの第1GOPのためにエンコードすると、開始編集ポイントでの可視復号化人為構造は排除される。10m 0sから10m 1sまでの元のエンコードに使用されたのと同じ量子化レベルを使用する最後の部分のエンコードは、図33(A) - 図34に関して前述されるのと同じであり、その結果生じるエンコードビデオは図35(B)にブロック図で示される。GOP閉鎖モードは、第6.

3.8項でのMPEG規格に説明される。ピクチャのグループの閉鎖モードフラグは1にセットされ、エンコードされたB-フレームが、逆方向予測だけを使用してエンコードされ、以前のGOPの基準フレームが使用されていないことを示す。しかし、双方向フレームに通常使用される基準フレームの内の1つが使用できないため、引き下げられた品質を排除するために手順が実行されない限り、B-フレームは、品質を劣化させてしまうだろう。

【0229】GOP閉鎖モードでのB-フレームの品質劣化の問題を解決するために、単一方向だけの予測が設定されるB-フレームのビット数は、フレームの品質が大幅に引き下げられないように増加される。B-フレームには単一方向だけの予測が設定されているので、ビット数は、ほぼ通常のP-フレームに使用されるビット数に増加するのが好ましい。ただし、ビット数は、希望される画質に応じて変化する。

【0230】図35(B)のエンコードされたビデオは、図35(C)でブロック図で示されるように、既にエンコードされたビデオに直接代入できる。

【0231】GOP閉鎖モードを使用するプロセスを説明するフローチャートは、図36に示される。図36では、ビデオはステップ800での圧縮フォーマットにエンコードされ、エンコードされたビデオに代入される必要があるビデオの新しいセクションは、ステップ780と782に関して記述されるのと類似した方法でステップ802で決定される。次に、ステップ804では、必要に応じて、エンコードされていないオリジナルビデオのGOP(1つまたは2つのGOP)を、ビデオの新しいセクションの立ち下がり端に付加する。

【0232】それから、ステップ786でビデオをエンコードする。新しいビデオの第1GOPは、ピクチャのグループの閉鎖モードでエンコードされ、基準フレームを見失った双方向に予測されるフレームのビット数が、ほぼ単一方向予想フレームであるP-フレームに使用されるビット数に増加される。エンコードされたビデオのビット総数が同じまとなる場合、二方向フレームに必要な余分なビットは他のフレームから削除されなければならない。これは、空間手動編集または時間的手動編集、あるいはその両方についての項に前述するように達成される。次に、ビデオの新しいセクションの残りは普通にエンコードされる。最後に、新規ビデオの立ち下がり端で付加された期間は、編集セグメントの終了ポイントでの復号化人為構造を防ぐために、追加された期間の元のエンコードの間に使用されたのと同じ量子化値を使用して、エンコードされる。最後に、新規にエンコードされたビデオが、ステップ808で最初にエンコードされたビデオに代入される。

【0233】前述のプロセス、および図34と36にブロック図されたプロセスのいずれかによって、代わりのエンコードされたセクションを含むエンコードされたビデオ

の最終バージョンでは、復号化人為構造の数が低減されるため、可視ビデオグリッチは減少し、復号化人為構造を防ぐためになんの訂正処置も講じられない場合と比較してビデオの品質が向上する。前記を説明する例は、GOPの境界での編集ポイントに関係する。ただし、前記教示は、GOP境界にない編集にも適用できる。

【0234】復号化人為構造を低減するプロセスは、再生VTR51およびビデオエンコーダ50とともに、ワークステーション10、30、および40を使用して実行される。しかし、それ以外のハードウェア代替策も可能であり、発明の前述した面は、プロセスを実行するために、任意の汎用コンピュータまたは特定のハードウェアを使用して実現できる。編集ポイントを決定するため、およびエンコードされるべきビデオの新規セクションの最後に付加されるビデオを決定するために使用される計算は、ワークステーション10またはワークステーション30のどちらかで実行される。

IX. エンコードに使用されるレート量子化モデリング

エンコードプロセスから生じるマクロブロックの列を表すのに必要とされるビット数を正確に概算するために、フレームは、フレームの標準量子化レベルを表す複数の量子化レベルを使用して事前にエンコードできる。それから、フレームは、他の量子化レベルだけではなく、代表的な量子化レベルを使用して実際にエンコードできる。図1Aに示される実施例では、ワークステーション10は事前に決定された量子化レベルをビデオエンコード装置50にロードする。それから、ビデオエンコード装置50は、事前エンコードを実行する。量子化モデルを作成するもっとも正確な方法では、可能な量子化レベルごとに各フレームをエンコードし、結果として生じる使用ビット数に関する統計を記録することが必要となる。しかし、このプロセスには、未処理ビデオでの n 回のパスが必要となる。この場合、 n は未処理ビデオのエンコードに使用される量子化レベルの総数である。この方法はひどく長いので、本発明は、従来の方法に要した時間の少くして最適の方法に類似した結果を作成する改善された方法となる。方法には、量子化レベルに対する代表的なビットレートを、事前エンコードフェーズの間に複数のフレームのそれぞれに計算できるように、複数のフレームのさまざまなパーツに異なった量子化レベルを割り当てることが必要になる。これにより、正確なビットレート対量子化レベルの統計を作成しながらも、各フレームの $n-1$ の事前エンコードが節約される。それから、これらの統計を補間し、そのビットレート対量子化レベルが事前エンコードプロセスで決定されなかったマクロブロックによって使用されるビット数を概算する。2つの代替実施例では、方法は未処理ビデオまたは未処理ビデオのすべてのフレームから過半数のフレームを事前エンコードする。

【0235】図37(A)には、2ポイントビットレート対量子化レベル線が複数のフレームに作成され、結果が複数のフレームに記憶されるように、事前エンコードプロセス内でマクロブロックに基づいて2つの異なった量子化レベルをマクロブロックに割り当てられる未処理ビデオのフレームが示される。代替例では、量子化レベルは、マクログループのグループまたはストリップで割り当てられる。図37(B)では、量子化レベルが、位置に基づく統計上の異常を回避するために、フレーム内のさまざまな場所に割り当てられるフレームでの4つの量子化レベルの均一な分配を説明する。図示されたフレームでは、各行の j 、行内の最初の量子化レベルは、 $qj \bmod 4$ である。この結果、4ポイントビットレート対量子化レベルの曲線が生じる。図37(C)では、図37(B)と同じプロセスが示されるが、量子化レベルはブロック様式で割り当てられている。このプロセスのシングルパスは正確なビットレート対量子化レベル曲線を作り出すが、代替実施例では、このプロセスは、量子化レベルごとに各フレームを分析することによって収集される統計に近づくさらに正確な統計を作成するために、複数のフレームで何度も繰り返される。

【0236】図38では、複数のフレームの量子化レベル対ビットレート特性を決定する一般的なプロセスを説明する。プロセスはステップ850で開始し、事前エンコードされる未処理ビデオを表すビデオ信号を獲得することにより、ステップ852で続行する。ステップ854では、プロセスは、第1フレーム、 i がフレームインデックスであるところの $i=1$ が事前エンコードされなければならないことを示す。ステップ856に続けると、フレーム $i=1$ の量子化レベルは、統計学上の異常を防止するためにフレームで均一に割り当てられる。ステップ858では、フレーム $i=1$ は、ステップ856で割り当てられる量子化レベルを使用して事前エンコードされる。ステップ858の結果は、ステップ856で割り当てられる量子化レベルごとのビットレートがフレーム $i=1$ に記録されるように、ステップ860で求められる。ステップ862では、方法が、複数のフレームのすべてが事前にエンコードされているかどうかを判断する。事前にエンコードされていたフレームは1つだけであったので、プロセスは、フレームインデックス i に1を加算してエンコードされるべきフレームとして第2フレームを指定するステップ866まで続く。それから、制御は、量子化レベルをフレーム $i=2$ に割り当てるステップ856に戻される。ステップ856、858、860、および862で割り当て、事前にエンコードし、ビットレートを決定するプロセスは、すべての残りのフレームについて続行される。ステップ862で、複数のフレームのすべてが事前にエンコードされていたと判断される場合、プロセスはステップ864に移動することにより終了する。ステップ864に到達すると、ビットレート対量子化レベル曲線は、ステップ856で割り当てられる量子化

レベルおよびその対応するビットレートを使用することにより、近似される。

【0237】未処理ビデオの複数のフレームにビットレート対量子化レベル曲線を作成したので、事前エンコードプロセスは、未処理ビデオの一部がデジタル記憶装置媒体、つまりコンパクトディスクに適合するようにエンコードされるエンコードプロセスに移動する。未処理ビデオセグメントの1つ、複数、過半数またはすべてが、この第2フェーズでエンコードされる。未処理ビデオセグメントを自動的にエンコードする必要がある場合は、エンコーダは、セグメント内の各フレームのマクロブロックに割り当てられる量子化レベルを決定するために、対応するビットレート対量子化レベル曲線と結び付いた各マクロブロックの活動基準を使用する。それから、セグメントの各フレームがエンコードされ、その結果として生じるビット数は、割り当てられた量子化レベルに基づいてビット概算数と比較される。結果として生じるビット数がビット予測数から大きく異なる場合、自動エンコーダは、エンコードされるマクロブロックの次の集合の量子化レベルを変更することによって、変動を補償しようと試みる。この連続プロセスにより、自動エンコーダはターゲットビット数に一致するように未処理ビデオ全体のセグメントをエンコードする。

【0238】さらに、ビデオシーケンスのもっと後のフレームのビットレート対量子化レベル曲線は、結果として生じるエンコードされたデジタルビデオを通して品質を調整する場合に使用できる。図39に図示されるように、フレームNは同じ量子化レベルにフレームMより少ないビットを作成すると知られている。したがって、フレームMは、フレームN内と同じ品質の画像を作成するためにはさらに多くのビットを必要とする。フレームNがそのすべての割り当てられたビットを使用しないで希望の品質でエンコードされると、エンコーダは、同じ品質のために、ユニット時間あたりさらに多くのビットを必要とするフレームMをエンコードする際に使用される余分なビットを前方に運ぶことができる。この機能は、M+N>1の場合にフレーム先読みバッファが1つしかないシステムでは利用できない。

X. データ構造を使用して、オーディオ、ビデオ、およびサブピクチャのデジタル信号を結合する
本特許明細書の前記部分は、おもに、エンコード圧縮デジタルオーディオ映像信号を作成するために、オーディオ映像情報のエンコードの間に使用される技法に関する。しかし、オーディオ情報およびビデオ情報が結合、記憶、およびテレビの上などで、オーディオ映像情報を作成し直す、または構築し直すために後で復号化されるためには、オーディオ情報、ビデオ情報、およびそれ以外の記憶され、後で復号化できるようにエンコードされていた情報を結合し、インタリーブする必要がある。オーディオ情報、視覚情報、およびそれ以外の情報の結合

は、図1Aに図示されるワークステーション20で発生するフォーマットプロセスとして、前記に簡略に説明した。フォーマットプロセスならびに結果として生じるデータ構造および信号について、ここで詳しく説明する。

【0239】ここで図1Aを見ると、MPEGビデオおよびエンコードされたオーディオはディスク21に記憶される。さらに、ビデオの上に表示またはオーバーレイあるいはその両方を行われる字幕または図形表記、あるいはその両方を含むサブピクチャ情報は、例えば、別個のワークステーションにより初期にエンコードされ、エンコードされたサブピクチャ情報は、フロッピーディスク、テープまたはその他の種類のディスクのようなデジタル記憶装置媒体に記憶される。この記憶されたサブピクチャ情報は、デジタル記憶装置22の内の1つにより読み取られ、フォーマットプロセスでディスク21内のエンコードされたオーディオとビデオと結合され、ワークステーション20により1つのデータストリームにフォーマットされる。フォーマットプロセスの出力は、ディスク21に記憶されてから、デジタル記憶装置22のデジタルテープに書き込まれる。それから、テープは既知の光ディスク製造法により光ディスクを作成するために使用される。したがって、本発明は、オーディオおよび視覚データのエンコードに関係するだけではなく、エンコードデータ、オーディオデータ、ビデオデータおよびサブピクチャデータを1つのデータストリーム、デジタル記憶装置媒体およびフォーマットされたオーディオ情報、ビデオ情報、およびサブピクチャ情報を記憶するその上でのデータ構造に結合するためのプロセスおよびシステムに関し、復号化プロセスおよびユーザまたは消費者に表示するためにオーディオ情報、ビデオ情報およびサブピクチャ情報を分離し、再構築するデコーダにも関する。光ディスク上またはフォーマットされたデータ内でのデータ構造が、メモリ内に記憶される情報の物理的な編成を分け与える特殊電子構造要素に関係することを注記すべきである。これらの特殊電子構造要素は、ビデオオンデマンドシステムに使用できるまたはデジタルビデオディスクプレーヤを使用してユーザまたは消費者のために復号化される光ディスク内に記憶される。

【0240】エンコードされたオーディオ映像情報用のディスクまたはそれ以外の記憶装置媒体のフォーマット構造の特定な特徴を説明する前に、本発明により改善される既知の規格について説明する。ISO/IEC 13818-1は、そこに引用される参考とともに、参照により本明細書に取り入れられるMPEGビデオおよびオーディオのシステム面を記述する。これ以降MPEGシステム説明と呼ばれるISO/IEC 13818-1は、ビデオストリームおよびオーディオストリームの基本的な多重化アプローチを説明する。これらのストリームは、パケット化基本ストリーム(PES)パケットと呼ばれるパケットに分割される。また、MPEGシステム説明には、1つ以上の独立したタイム

ベースのある1つ以上のプログラムを1つのストリームに結合する移送ストリームの使用についても説明される。MPEGシステム説明は、複数の基本ストリームを同期させるために使用されるプレゼンテーション時刻記録 (PTS) の使用について説明する。時刻記録は、一般的には90kHzという単位で、システムクロックリファレンス (SCR)、プログラムクロックリファレンス (PCR)、およびオプションの基本ストリームクロックリファレンス (ESCR) といっしょに使用される。データがMPEGシステム説明に従い、どのようにフォーマットされるのかについての詳細の完全な説明は、ISO/IEC 13818-1に記載される。

【0241】ISO/IEC 13818-1に基づいたフォーマットプロセスを使用するのに加えて、データは、また、ISO 9660, 1988、情報処理—ボリューム、および本明細書に参考により取り入れられる情報交換用CD-ROMのファイル構造に従ってフォーマットおよび記憶される。この規格は、ディスクボリューム（音量ではなく、ディスクの内容）およびファイル構造がどのように編成されるのかを記述する。

【0242】オーディオデータ、ビデオデータ、およびサブピクチャデータを記憶するために使用される光ディスクの現在の時刻での実施例は、片面あたり5ギガバイトの容量で、総記憶容量10ギガバイトの単層両面ディスクである。将来のフォーマットでは、記憶容量を増加するために多層ディスクを見越し、読み取り専用ディスクの使用に加えて、追記技術および上書き技術も使用できる。本発明とともに使用できるディスクのそれ以外の面については、1994年9月13日に提出された「光ディスクおよび光ディスク装置」という米国特許明細書08/304,848に説明する。ディスクの外径は120 mmであるのが好ましい。

【0243】図40では、究極的には光ディスクに具備されるワークステーション20のフォーマットされた出力の構造を説明する。この構造900は、リードイン領域902を含む。リードイン領域にはディスクから読み取られるデータのそれ以降の処理を制御する個々のフィールドが含まれる。リードイン領域900に記憶される例示的な引込み情報は、本明細書に参考により取り入れられるクックソン (Cookson) その他の米国特許5,400,077の図3にブロック図で示される。

【0244】システム領域904およびボリューム管理情報906はISO 9660に準拠する。ボリューム管理情報906には、主要なボリューム記述子、ボリューム記述子設定終止プログラム、ディレクトリレコード、およびパステブルレコードが含まれる。主要なボリューム記述子には、ディスクがビデオディスクであるのか、またはオーディオディスクであるのかを示す記述子型、ディスクがどの規格に準拠しているのかを示す標準識別子、システム識別子、ボリューム識別子、ボリュームサイズ、論理ブロックサイズ、パステブルサイズなどのディスクボ

リュームの基本的な情報が記憶される。システムのこれらのフィールドおよびそれ以外のフィールド、およびボリューム管理情報は、ISO 9660に詳説されているので、その説明は簡略さを期すために省略される。さらに、希望する場合には、ISO 9660に準拠する補足ボリューム記述子もある。

【0245】ボリューム記述子設定終止プログラムには、ボリューム記述子型、標準識別子、およびボリューム記述子バージョンが記憶される。ディレクトリレコードには、ISO 9660に従ったディスクのさまざまなディレクトリ情報が記憶される。パステブルレコードには、ISO 9660に記述されるように、L型パステブルおよびM型パステブルが記憶される。

【0246】ディスク情報ファイル908は、ディスクの内容についてのさらなる情報および光ディスクをプレイするときにユーザが選択できるメニュー情報に記憶される。ディスク情報ファイル908の詳細は、図41に関して説明される。

【0247】各ディスクには、最大99データファイルまで最小1データファイル910が格納される。例えば、ディスク上に2つの異なったオーディオ映像プログラムが記憶されている場合、ディスク情報ファイル908には、ユーザが使用を希望するデータファイルのどれかをユーザが選択するためのメニューを作成するのに必要な情報が記憶される。データファイルは、図42-56に関連して詳説される。リードアウト領域16は、ディスク、処理がどのように実行されるのかを記述するそれ以外の情報、または関連するディスクに関する情報が記憶されることがある。

【0248】図41は、図40のディスク情報ファイル908を説明する。ディスク情報ファイル908には、ファイル管理情報920およびユーザが対話するメニューを構成する情報であるメニューデータ922が格納される。メニューデータは、ビデオデータ、オーディオデータ、およびサブピクチャデータを含む。ファイル管理情報920は、ファイル管理テーブル924、ディスク構造情報926、メニュー構造情報928、および複数メニューセル情報フィールド932と934が含まれるメニューセル情報テーブル930を格納する。

【0249】ファイル管理テーブル924には、ディスク情報ファイルのあらゆる面に関する情報が含まれる。ファイル管理テーブル924内には、ファイル名を識別するためのボリューム管理情報906の対応するディレクトリレコード内の対応するファイル識別子と同じ内容を記述するファイル名が記憶される。論理ブロックの数を使用してファイルのサイズを記述するフィールドがある。本明細書全体で、「論理ブロック」という用語は、ISO 9660に定義されるように使用される。論理ブロック数を使用するファイル管理テーブルのサイズ、ファイルの先頭からの相対論理ブロック番号を使用するディスク構造情

報926の開始アドレス、ファイルの先頭からの相対論理ブロック番号を使用して記述されるメニュー構造情報928の開始アドレス、ファイルの先頭からの相対論理ブロック番号を使用するメニューセル情報テーブル930の開始アドレスを記述するエントリがある。メニューデータがない場合は、このフィールドの値はゼロになる。ファイルの先頭からの相対論理ブロック番号を使用するメニューデータ922の開始アドレスを記述するフィールドがある。メニューデータが存在しない場合、このエントリはゼロである。

【0250】ファイル管理テーブル924には、メニューデータ用のビデオのビデオ属性を記述するフィールドもある。この情報は、ビデオ圧縮モードがMPEG-1であったのか、MPEG-2であったのか、ビデオのフレームレート（毎秒29.97フレームまたは毎秒25フレームのどちらか）、表示アスペクト比が3/4であるか、または9/16であるのかを示すフィールド、およびパンスキャンとレターボックスモードの両方が許可されていることを示すか、パンスキャンモードが許可されているが、レターボックスモードが禁止されていることを示すか、あるいはパンスキャンモードが禁止され、レターボックスモードが許可されていることを示すディスプレイモードフィールドを含む。

【0251】ビデオ属性を記述するフィールドと同様に、ファイル管理テーブル924のオーディオストリームを記述するフィールドもある。このフィールドには、オーディオがドルビーAC-3に従ってエンコードされたのかどうか、オーディオがMPEGオーディオであるのかどうか、あるいはオーディオが線形PCMオーディオ（48 kHzで16ビット）を含むオーディオコーディングモードの表示が指定される。また、オーディオがモノラル、ステレオ、またはデジタルサラウンドであるのかを含むオーディオコーディングモードの表示も指定される。

【0252】ファイル管理テーブル924には、さらに、メニューデータ922のサブピクチャストリーム属性を記述するフィールドも入る。このフィールドは、以下に説明するように、ランレングスサブピクチャコーディングモードがサブピクチャに使用されることを示す。フィールド管理テーブル924には、メニューのためのすべてのサブピクチャに使用される16セットのカラーパレットを記述するフィールドも入る。パレット番号1から16には、輝度信号Y、カラー差異信号Cr=R-Y、およびカラー差異信号Cb=B-Yが指定される。

【0253】ディスク構造情報926には、ディスクに記憶されるビデオファイルおよびオーディオファイルの構成情報が含まれる。ディスク構造情報926には、ファイル名を識別するために使用されるディレクトリレコード内での対応するファイル識別子と同じ内容を記述するファイル名が含まれる。ファイルをディスク構造情報ファイルとして識別するファイル識別子フィールドがある。

図40に示されるデータファイル910および914のようなディスク上のデータファイルの数を示すフィールドがある。また、ファイルにビデオ情報とオーディオ情報の両方が記憶されるのか、あるいはオーディオ情報だけが記憶されるのか、およびファイルがカラオケファイルであるかどうかを含むデータファイルのそれぞれのタイプを示すフィールドもある。

【0254】ディスク構造情報926には、データファイル内のサブピクチャとオーディオ情報を含むフィールドも含まれる。このフィールド内には、サブピクチャチャンネル数も示される。各サブピクチャチャンネルは、異なったサブピクチャ言語のようなさまざまなサブピクチャ情報を表示する場合に使用できる。また、オーディオストリーム数の表示、チャンネル番号順の各サブピクチャチャンネルの言語コードの表示、オーディオストリーム番号順のオーディオストリームの言語コードの連続記述、およびファイルのパレンタル管理とアングル情報の記述もある。パレンタル管理情報は、表示される情報のある特定のシーケンスを記述する場合に使用される。パレンタル管理情報は、シーケンス内に含まれる特定のセルの最大親レベルである。レベルは、1から5で変化し、レベルが高いほど、見聞きされる内容に対する制限が多くなる。シーケンスとセル、およびパレンタル管理スキームの詳細は、以下に説明する。ビデオシーケンスのアングルの数を記述するフィールドもある。例えば、ミュージックビデオに、ミュージックビデオ全体を通して3つの異なったカメラビューがある場合、ユーザはどの特定のカメラアングルを表示に希望するのかを選択できる。また、シーケンス内のプログラム数を示すフィールドもある。

【0255】メニュー構造情報928には、各メニュータイプの開始セル番号が含まれる。ディスプレイシーケンスには複数のセルがある。例えば、1つのシーケンスは最高256のセルを格納できる。セルは、後述する再生情報（PBI）パックと次のPBIの直前のパックの間のGOP（ピクチャのグループ）の整数である。ビデオ用データは、データを連続して再生する必要があるレコード単位としてセルの中に分離される。セルには、表示のための情報が含まれ、その目的に応じて分割される。セル内の第1オーディオパックとおよび第1サブピクチャパックが、セルの先頭にあるPBIパックの後ろに記録されるビデオヘッドパック中のPTSに近いプレゼンテーション時刻記録（PTS）を含むのが望ましい。異なったセルは、希望されるアングルおよびプログラムの親制約に応じて表示される。例えば、より若年層の視聴者向けには不適切であった映画のあるセクションがあったと仮定する。表示できるパレンタル管理レベルに応じて、2つのシーケンス共用セルが存在する。例えば、制約されていない表示の場合、ある特定のシーケンスにセル1、2、3、および5を表示できる。制約レベルの表示の場合は、そのシーケンスに表示されるセルは1、2、4、および5である場

合がある。

【0256】異なったアングルに関しては、同時進行し、その再生に要する時間が実質上等しい、複数のカメラアングルで各ビデオを記録させることができる。各アングルは、アングルセルと呼ばれる1つのセルから構成される。同時に先行する複数のアングルセルの集合が、アングルブロックと呼ばれる。シーケンス内の複数オロケーションでアングルブロックが限定される場合、各アングル番号のアングルセル数およびカメラアングルは、アングルブロックを通して同じでなければならない。

【0257】メニューセルとは、ユーザに、ボリューム内の各ビデオファイルまたはオーディオファイルの内容を知らせるメニュー画面を構成するセルである。メニューセルの集合は、メニュー用データとしてディスク情報ファイルとともに記録される。メニュータイプと呼ばれる、メニューセルタイプは、タイトルメニュー、プログラムメニュー、オーディオメニュー、サブピクチャメニュー、およびアングルメニューに分類される。メニュー画面には、メニュー用ビデオが備えられる。再生装置の画面上に表示される選択用アイテムは、メニューセルのサブピクチャデータとして記録される。

【0258】メニューセル情報テーブル930は、図41の932と934のような複数のメニューセル情報が記憶されるテーブルである。1つのメニュー画面に1つのメニューセルが具備される。メニューセル情報は、複数のフィールドを含む。第1フィールドは、ディスクの情報のコピーが許可されているかどうか、パレンタル管理のレベル、メニューがタイトルメニューであるかどうかなどのメニューセルタイプ、プログラムメニュー、オーディオメニュー、サブピクチャメニュー、またはアングルメニューおよびメニューの言語コードを示す。各言語にはコードが割り当てられ、言語コードサブフィールド内に現れる特殊コードが、メニューと使用される言語を示す。

【0259】メニュー画面上のサブピクチャによって表示される選択アイテムの開始番号および選択アイテム数を記述するフィールドもある。選択アイテム開始番号は、メニュー画面上の選択アイテムの最小数を記述し、メニューの次のページが存在するかどうかを示すフィールド、および選択できる1と16の間のアイテム数を記述する番号がある。メニューセル情報は、メニューセルの開始パックのシステムクロックリファレンス(SCR)の下位32ビット、メニューセルの開始アドレスをファイルの先頭からの相対論理番号で記述するフィールド、およびメニューセルを構築する論理ブロック数を記述するフィールドも含む。

【0260】メニューデータ922は、メニューのために表示される実際の情報を含む。この情報は、図40のデータファイル910と914の情報が構築されるのと同じように構築され、したがって各面の綿密な記述は冗長な記述を防ぐために省略される。メニュー画面で選択されるアイ

テムは、サブピクチャ情報で形成される。ディスクプレーヤによるメニュー選択の処理は、コンピュータコードを実行するマイクロプロセッサを使用して実行される。タイトルのためのメニュー選択が行われると、選択の結果はファイルの最初または選択されたタイトル番号に対応するシーケンス番号から再生される。プログラムメニュー選択の場合、選択されたプログラムは選択されたプログラムの最初から再生される。メニューからのオーディオ選択の場合、オーディオストリームは選択されたオーディオストリームに対応するように設定される。同様に、サブピクチャメニュー選択の場合、チャンネル番号は選択されたサブピクチャチャンネルに等しく設定される。最後に、メニューからのアングル選択は、アングルセル番号を選択された番号に等しくなるように変更する。それ以外の情報は、希望に応じて含まれ、メニューを通して選択できる。

【0261】図40のデータファイル910のようなデータファイルには、図42にブロック図で示されるような2つのセクション、つまりデータファイル管理情報940およびオーディオ情報、ビデオ情報およびサブピクチャ情報を表示するために使用される実際のデータ942が記憶される。データファイル管理情報940には、4つの主要な構成要素、つまりデータファイル管理テーブル944、シーケンス情報テーブル946、セル情報テーブル948およびデータ検索マップ950を含む。

【0262】データファイル管理テーブル944は、データファイル910の各面に関する情報を含む。この情報には、ファイル名を識別するためのディレクトリレコード内の対応するファイル識別子と同じ内容であるファイル名のフィールド、オーディオ情報とビデオ情報の両方がファイルに記憶されることを示すファイル識別子、論理ブロック数により定義されるファイルのサイズ、ファイル管理テーブルのサイズ、ファイルのシーケンス番号、ファイルのセルの数、ディスク構造情報(DSI)パックの数、シーケンス情報テーブル946の開始アドレス、セル情報テーブル948の開始アドレス、データ検索マップの開始アドレス、データ942の開始アドレス、MPEG-1またはMPEG-2のようなビデオ圧縮モードを含むデータ942のビデオ属性を記述するフィールド、レートが毎秒29.97フレームであるのか、毎秒25フレームであるのかを含むフレームレート、表示アスペクト比が3/4であるのか、9/16であるのか、またはディスプレイモードによってパンスキャンフォーマットとレターボックスフォーマットの1つが可能とされるのか、両方が可能とされるのかを示す表示アスペクト比を含む。

【0263】データファイル管理情報は、さらに、ファイル内のオーディオストリーム数を、オーディオコーディングモードのようなオーディオストリーム属性、モノラル、ステレオ、またはデジタルサラウンドのようなオーディオモード、特殊言語が含まれているのかどうかを

示すオーディオタイプを記述するフィールド、および言語のコード番号を示す特殊コードのフィールドも含む。

【0264】データファイル管理テーブル944は、さらに、サブピクチャチャンネル数を記述するフィールド、サブピクチャのランレングスコーディングが使用されること、および指定言語および指定される言語があるかどうかを示すフィールドのようなサブピクチャチャンネル属性を記述するフィールドを含む。さらに、ファイルのすべてのサブピクチャチャンネルで使用される16のカラーパレットのY、Cr、およびCbカラーを示すフィールドがある。また、シーケンス情報テーブルのサイズ、最小セル番号と後続の連続セルの数を記述するフィールド、およびファイルの1つのシーケンス内で使用されるセル番号の範囲を記述するフィールドもある。1つのシーケンス内で使用されるセル番号は、32の連続番号で1つの集合を作り、最小セル番号を集合の最初の番号となるように格納することにより計算される数の集合に入れられなければならない。このフィールドには、シーケンスで 사용되는セル数、およびシーケンスで 사용되는最小セル数を含むサブフィールドがある。

【0265】シーケンス情報テーブル946には、複数のシーケンス情報エントリ952と954を含む。シーケンスとは、このシーケンスにより指定される範囲内のセルが選択式で再生される順序のことである。完全および接続という2種類のシーケンスがある。完全型シーケンスは、それ自体を再生した後に終了する。接続型シーケンスは、連続して再生される。参照番号952により示されるシーケンス情報および954により示されるシーケンス情報iのようなシーケンス情報番号は、シーケンス番号に等しく、1で開始するシーケンス情報テーブルi、記述順序で番号が付けられる。シーケンス情報テーブルのシーケンス情報エントリは、完全型シーケンスまたは接続型シーケンスの順序で記述されている。各シーケンス情報テーブルは、完了型シーケンスまたは接続型シーケンスの順序で記述される。各シーケンス情報には、シーケンスタイプ、格納されるプログラム数、格納されるセル数、シーケンス再生時間、接続型シーケンス数、接続可能シーケンス番号、およびシーケンス制御情報を示すフィールドを示すフィールドを含む。

【0266】シーケンス情報のシーケンスタイプフィールドは、シーケンスのコピーおよびパレンタル管理を記述する。コピーが許可されているのか、または禁止されているのかを示すサブフィールド、パレンタル管理のレベルを1-5で指定するサブフィールド、完了型シーケンス、接続型ヘッドシーケンス、接続型中間シーケンスまたは接続タイプとシーケンス内の1つである可能性があるシーケンスタイプを示すサブフィールドがある。また、シーケンスが、カラオケアプリケーションで使用するかどうかを示すアプリケーション型サブフィールドもある。

【0267】格納プログラム数フィールドは、最高1つのシーケンス内に100というプログラム数を記述する。格納セル数フィールドは、ある特定のシーケンス内での、合計256セルまでのセル総数を記述する。シーケンスまたは再生時間は、時、分、秒およびビデオフレームによる単位でこのシーケンスの総再生時間を記述する。

【0268】接続型シーケンス数フィールドには、最大数が8であるその直後に接続可能なシーケンス数を指定する。接続可能シーケンス番号フィールドは、接続可能シーケンスのシーケンス番号と指定される数のシーケンスのパレンタル管理のレベルを記述する。接続可能シーケンスの選択番号は、記述された順序で1から割り当てられる。このフィールドには、パレンタル管理番号のサブフィールド、およびシーケンス番号のサブフィールドが指定される。シーケンス情報の最後のフィールドは、再生順でシーケンスに含まれるセルを記述するシーケンス制御情報である。このフィールドも、複数のセルの集合であるアングルブロックの記述に使用される。それらのセルの再生時間は、実質上等しい。各再生セクションは、1つのセルで構成される。アングルブロックは、多くても9個のセルを含み、第1セルに1という番号が付けられる。シーケンス制御情報のサブフィールドには、実質上再生される複数のセルを結合する再生単位としてプログラムを定義するプログラム番号を指定する。以下のセルに連続して進むのか、休止するのか、あるいはシーケンスの最後に到達したかを示すセル再生制御のサブフィールドがあり、ブロックが構成セルではないかどうか、ブロック構成の第1セルがあるかどうか、それがブロック構成内のセルなのか、ブロック構成内の最後のセルであるのかを示すブロックモードサブフィールドがある。ブロックがないかどうか、またはアングルブロックがあるかどうかを示すブロック型サブフィールドがある。最後に、再生対象のセルのセル番号を示すサブフィールドがある。

【0269】データファイル管理情報940は、さらに、複数のセル情報エントリ956と958を含むセル情報テーブル948を含む。各セル情報エントリは、コピーが許可されているのか、禁止されているのかを示すセルタイプ情報、およびパレンタル管理レベルを含む。また、時、分、秒およびビデオフレームによりセルのセル再生総時間を記述するセル再生時間を示すフィールドもある。セルの第1バックに記述されるシステムクロックリファレンスの下位32ビットを記述するフィールド、ファイルの始めからの相対論理ブロック番号でセルの開始ポイント得のアドレスを記述するフィールド、およびセル内に含まれる論理ブロック数を記述するフィールドがある。

【0270】ファイル管理情報940のデータの最後のセクションは、データ検索マップ950である。データ検索マップは、データ検索情報パック960と962の複数のポイントを備える。再生情報(PBI)パックポイントと呼ばれ

るこれらのポイントは、データ942内に存在する再生情報のアドレスを指定する。高速順方向モードまたは高速反転モードを実行し、情報の表示を可能にするためには、MPEG-2ビデオシーケンスのイントラピクチャに依存するのが最適です。これらのイントラピクチャは、そのアドレスがデータ検索マップ内に記憶されるデータ942内の再生情報パックを使用して配置される。PBIパックの詳説は、以下で行う。

【0271】図43にブロック図で示されるデータファイル910のデータ942は、再生情報(PBI)、ビデオ情報、サブピクチャ情報、およびオーディオ情報のインタリーブされたパックを含む。例えば、図43では、データは再生情報940、ビデオ情報942と944、サブピクチャ情報946、オーディオ情報948などを含む。データ942内の情報の構造、およびプレゼンテーションタイミングは、MPEGシステム説明(ISO/IEC 13818-1)に定義されるプログラムストリームに準拠する。しかし、再生情報およびサブピクチャ情報、ならびにこの情報が、記憶、エンコード、および復号化される方法が、本発明の一部を構成する。

【0272】図43でデータ942から構成されるさまざまな情報は、情報パックを含む。例示的なパック970Aおよび980Aが、図44Aと44Bに示される。各パックは、パックヘッダ972Aまたは972B、およびビデオ情報、オーディオ情報、サブピクチャ情報、または再生情報から構成される。

【0273】パック970Aの構造は、パケット982Aが2034バイトと2027バイトの間で占有するときに使用される。スタッフィングフィールド980Aによって、1から8バイトのスタッフィングがパック総サイズを2048バイトにできるようにする。ビデオ情報、オーディオ情報、サブピクチャ情報または再生情報が2027バイトを下回るとき、図44Bに示されるパック構造970Bが使用され、1バイトのスタッフィング970Bおよびパケット982Bと984Bのビット総数を2034バイトにあるようにする埋め込み984のパケットがある。パケットヘッダ972Aと972Bのそれぞれには、パック開始フィールド974、システムクロックリファレンス976、多重化(MUX)レートフィールド978、およびスタッフィングフィールド980がある。図44Aと44Bのパケットは、ISO/IEC13818に従って構築され、プログラムストリームパック内のフィールドのセマンティック定義に関するその第2.5.3.4項に説明される。

【0274】図45は、再生情報(PBI)パックを示している。パックヘッダ972は、図44Aと44Bに関して記述されるように、構築される。図45のシステムヘッダ986は、ISO/IEC13818-1に記述されるプログラムストリームシステムヘッダの要件に準じて構築される。システムヘッダ内のフィールドのセマンティック定義は、ISO/IEC 13818-1の第2.5.3.6項に説明される。

【0275】パケットヘッダ988は、ISO/IEC 13818-1の第2.4.3項に記述されるPESパケット内のフィールドのセ

マンティック定義に説明されるフィールドに従って構築される。ただし、パケットヘッダ988では、プレゼンテーション時刻記録までのフィールドだけが必要とされる。

【0276】再生情報パックは、さらに、サブストリームIDフィールド990を含む。これは、サブストリームの識別を示す8ビットのフィールドである。使用できる別の識別は、サブピクチャストリーム、ビデオ空白化情報(VBI)ストリーム、AC-3ストリーム、または線形PCMストリームである。MPEGオーディオサブストリームなどの他のストリームタイプを入れることも可能である。サブストリームID990はビデオ空白化情報992のデータを先行するので、サブストリームIDは、後続の情報がビデオ空白化情報ストリームであることを示すように設定される。ビデオ空白化情報992のデータは、ISO/IEC 13818-1に従って構築され、ISO/IEC 13818-1に必要となる意味論に準拠する。

【0277】図45に示される再生情報パックの中には、データ検索情報996のデータを格納するデータ検索情報パケットのパケットヘッダ994も示される。パケットヘッダ994には、24ビット値000001hが設定されるパケット開始コードプレフィックス、データ検索情報996が以前に定義された規格に準拠していないことを示すストリーム識別、パケット長を示すフィールドを含む。データ検索情報996はMPEG規格では開示されていないので、ストリームIDは、後続の情報がデータストリームの私的なタイプであることを示す。

【0278】図45を説明するデータ検索情報996のデータの特定の内容が図46に示される。データ検索情報996は、一般情報1000、アングルアドレス情報1020、ハイライト情報1022、同期再生情報1024、およびPBIパックアドレス情報1036を含む。

【0279】一般情報1000は、図47にさらに詳細にブロック図で示される。前記のように、データ検索情報の目的とは、デジタルエンコードされた情報の迅速な検索を可能にすることであり、特に、ビデオ復号化中に実行されるトリックモード用イントラピクチャオ高速位置発見を可能にする。したがって、図46に示される再生情報パックは、ピクチャのグループ(GOP)のそれぞれのヘッドパックであり、I-ピクチャで開始するビデオヘッドパックの直前に割り当てられる。一般情報1000は、I-ピクチャの再生開始時刻であるGOP1002の時刻記録である。次に、PBIのシステムクロックリファレンス(SCR)がある。これは、PBIパックヘッダに記述されるSCRの下位32ビットを記述する。1006は、I-ピクチャの最終アドレスを示し、ファイルの初めからの相対論理ブロック番号によりI-ピクチャの最終データが記録されるパックのアドレスを記述する。セル番号はフィールド1008に記憶され、GOPが属するセル番号を記述する。フィールド1010は、制約のレベルおよびコピーが許可されているかどうかも含

むパレンタル管理情報を記述する。最後に、フィールド1012は、ファイルのヘッドからの相対論理ブロック番号を使用するこのPBIパックのアドレスであるPBIパックアドレスを記述する。

【0280】図46に示されるデータ検索情報996のフィールド1020は、アングルアドレス情報である。アングルアドレス情報には、その他のアングルのアドレス情報が含まれる。アングルアドレス情報には、それぞれが異なったアングルセル番号に関する9つのフィールドが含まれる。アングルアドレス情報1020の各フィールドの場合、ある特定のPBIパックが属するセルがアングルブロックを構成する場合、アングルアドレス情報のフィールドは、対応するデータ検索情報のビデオ開始プレゼンテーション時刻記録を超えないもっとも近いビデオ開始プレゼンテーション時刻記録を持つアングルセル番号のPBIパックのアドレスを記述する。アングルセルが存在しない場合、またはアングルブロックが構成されない場合、特定のセルのアングルアドレス情報ははゼロに設定される。

【0281】データ検索情報996の第3セクションは、ハイライト情報1022である。ハイライト情報は、メニュー画面上での選択アイテムの位置、および選択時に変更されるその色とコントラストを記述する。この情報は、セルがメニューの一部である場合にだけ使用される。ハイライト情報には、選択アイテム開始番号およびメニュー画面上でサブピクチャにより表示されるアイテム数を記述するフィールドがある。選択アイテム開始番号およびアイテム数の特定のサブフィールドには、メニュー画面上での選択アイテムの最小数を記述するサブフィールド、メニューの次ページが存在するかどうかを記述するサブフィールド、およびメニュー画面上での選択アイテムの数を記述するサブフィールドがある。ハイライト情報の第2フィールドには、選択アイテムの位置、色、およびコントラストが含まれる。このフィールドは、メニュー画面上の選択アイテムごとの矩形表示領域、および選択時に変更される対応する色とコントラストを記述する。矩形表示領域は、X-Y座標面内に限定される。選択アイテムの位置、色、およびコントラストを記述するこのフィールド内の情報には、メニューの矩形領域を限定すること、およびさまざまなタイプのピクセルが、選択時にどのようにして変更されるのかを記述する情報を含む。さらに詳細に後述するように、サブピクチャは4種類の異なったピクセルを酌量する。テキストやそれ以外の情報を表示するために使用できる第1強調ピクセルと第2強調ピクセル、線描画またはその他の図形パターンのようなパターンを表示するために使用されるパターンピクセル、およびバックグラウンドピクセルがある。以上4種類の異なった情報がハイライトまたは選択される場合、アイテムがハイライトされるときに選択される色、およびハイライトされたときのピクセルのコントラ

ストも知っておく必要がある。したがって、4種類の異なったピクセルが選択されたときの、その色およびコントラストを記述するハイライト情報内に8つのサブフィールドがある。

【0282】データ検索情報996内の5番目のアイテムは同期再生情報1024である。同期再生情報1024は、図48に詳しくブロック図で示される。同期再生情報の目的とは、ビデオデータと同期するオーディオとサブピクチャの開始時間とアドレス情報を検出できるようにすることである。図48に示される同期再生情報1024には、オーディオのプレゼンテーション時刻記録のターゲットであるオーディオパックアドレスが含まれる。このフィールドの最上位ビットは、オーディオパックがこのPBIの前に位置しているのか、このPBIの後に位置しているのかを示す。オーディオ1028のプレゼンテーション時刻記録(PTS)は、ビデオ開始PTSからの相対プレゼンテーション時刻記録により、I-ピクチャの再生開始時間の直後に再生開始時間が設定されたオーディオパケットのPTSを記述する。オーディオのプレゼンテーション時刻記録は、プレゼンテーション時刻記録をそのために記憶させるために、最高8つのオーディオストリームを酌量する。さらに、GOPの間に再生されるオーディオフレームを構成するオーディオパケットがするかどうかを示すサブフィールドがある。ピクチャにオーディオがない場合には、オーディオのプレゼンテーション時刻記録を記憶したり、復号化する必要はない。最高8つのオーディオチャネルがあるので、各オーディオチャネルには対応するアドレスフィールドと時刻記録フィールドがある。

【0283】データ構造1024も、問題のサブピクチャパックの、PBIパックに相対したアドレスをを記述するサブピクチャパックアドレス1030を含む。サブピクチャのプレゼンテーション時刻記録は、対応するサブピクチャ情報の同期を記述する。GOP再生の間に再生される必要があるサブピクチャユニットが存在するかどうか、サブピクチャの再生開始時間がビデオプレゼンテーション時刻記録の前かどうか、GOP再生中に再生されるサブピクチャユニットの製作開始時間、およびGOP再生中に再生されるサブピクチャユニットの再生終了時間を示すサブフィールドがある。1つのデータストリームに含めることができる最高32のサブピクチャのように、存在するそれぞれのサブピクチャには、対応するアドレスと時刻記録のフィールドがある。

【0284】データ検索情報996の図46に示される最後のセクションは、PBIパックアドレス情報である。PBIパックアドレス情報は、ファイル内の他のPBIパックのアドレスを構成し、他のPBIパックの位置が迅速に見つけられるようにする。PBIパックアドレス情報は、相対論理ブロック番号を使用して周囲のPBIパックのアドレスを記述する。PBIパックアドレス情報には、それぞれ次のPBIパックアドレスと前のPBIパックアドレスを記述す

る2つのフィールドがある。現在のPBIパックアドレスからの相対論理ブロック番号を使用して、GOPのすぐ先にあるGOPに属するPBIパックのアドレスを、 $n \times 0.5$ 秒と $(n+1) \times 0.5$ 秒の間であるその累積した再生時間で定義し、複数のPBIパックがこの範囲に該当する場合は、 $n \times 0.5$ 秒、 $1 \leq b \leq 15$ を選択する15のフィールドがある。

【0285】現在のPBIパックアドレスからの相対論理ブロック番号を使用して、GOPに属するPBIパックのアドレスを、 $n \times 0.5$ 秒と $(n+1) \times 0.5$ 秒の間のその累積再生時間で定義し、複数のパックがこの範囲に該当する場合には、 $n \times 0.5$ 秒、 $-15 \leq n \leq -1$ にもっとも近いPBIパックを選択する15のフィールドがある。

【0286】現在のPBIパックアドレスからの相対論理ブロック番号を使用して、GOPのすぐ先のGOPに属するPBIパックアドレスを $n \times 0.5$ 秒と $[n \times 0.5 + 1]$ 秒の間のその累積再生時間で定義し、複数のPBIパックがこの範囲に該当する場合には、 $n \times 0.5$ 秒、 $n=20$ または $n=60$ にもっとも近いPBIパックを選択する2つのフィールドがある。

【0287】現在のPBIパックアドレスからの相対論理ブロック番号を使用して、GOPに属するPBIのアドレスを $n \times 0.5$ 秒と $[n \times 0.5 + 1]$ 秒の間のその累積再生時間で定義し、複数のPBIパックがこの範囲に該当する場合は、PBIパックを $n \times 0.5$ 秒、 $n=-20$ または $n=-60$ にもっとも近いPBIパックを選択する2つのフィールドがある。

【0288】ビデオパック1040の構造は図49にブロック図で示される。ビデオパックは、図44Aと44Bに示されるヘッダ972Aまたは972Bに従って構築されたパックヘッダ972を備える。ビデオ用のパケットヘッダ1046は、ビデオデータ1048のように、ISO/IEC 13818-1に従って構築される。ただし、MPEGビデオ信号として表記されるビデオデータ1048は、前記技法に従いエンコードするのが好ましい。

【0289】MPEGエンコードされたピクチャのグループは、図50の1050として示される。このGOP 1050を複数のビデオパック1040に分割し、ビデオストリーム1060を作成する方法は、図50の底部にブロック図で示される。ピクチャグループ1050は、ビデオパックに適合するために適切なユニットに分割される。最後のビデオパック1040G、ビデオパックjは、パックのそれぞれに使用される完全な2048バイトを使用しないので、最後のパックjを2048バイトに等しくするには、スタッフィングが使用される。

【0290】図51Aは、MPEGオーディオのビデオパックを示している。このビデオパックは、パックヘッダ972、MPEG要件に準拠して構築されるパケットヘッダ1084、およびエンコードされたMPEGオーディオデータ1086を含む。

【0291】図51Bは、線形PCMオーディオまたはドルビーAC-3オーディオのどちらかと準拠して構築されるオー

ディオパック1090を示している。図51Bにブロック図で示されるパケットの場合、パックヘッダ972は前記のように配列され、パケットヘッダ1094はMPEG要件に準拠して構築され、stream_idフィールドは線形PCMまたはAC-3のどちらかに対応する私的オーディオデータストリームを示す。サブストリームid 1096は、オーディオデータが線形PCM用であるか、AC-3（またはMPEG）用であるかを示す。オーディオフレームポインタ1098は、オーディオパケット内のオーディオフレームヘッダ数を示すフレームヘッダ数を具備し、ポインタの最後のバイトからの相対バイト数を使用して、オーディオフレームの第1アクセス装置の先頭を示す第1アクセス装置ポインタも備える。線形PCMオーディオにとって、各フレームがフレームヘッダとフレームデータを備えるオーディオフレームに分割されるのは普通のことである。オーディオフレームポインタ1098も同様にドルビーAC-3で使用できる。オーディオデータ1100は、対応するオーディオエンコード方法に従って構築される。

【0292】図52は、エンコードされたオーディオストリーム1100を示す。このオーディオストリームは、個々のパック1122を構成するオーディオパックのストリーム1120に変換される。各パックの長さは2048バイトで、最後のオーディオパック1122G、オーディオパックjは、長さを2048バイトにするために、スタッフィングビットを付加することによって調整される。

【0293】本発明の新規の特徴とは、サブピクチャの使用およびサブピクチャパックのオーディオ情報とビデオ情報とのインタリーブである。サブピクチャ情報を使用すると、グラフィック情報の表示が可能になり、MPEGビデオでのサブピクチャ情報の重複が酌量できる。本発明により使用されるサブピクチャ情報と従来のクローズキャプション字幕の間の大きな相違点とは、サブピクチャ情報がビットマップ化されたグラフィックであるのに対し、クローズキャプションでは、デコーダに記憶されている文字セット内の文字を表示するために文字符号が伝送されるという点である。したがって、従来のクローズキャプションは、デコーダ内の文字セットによって制限される。しかし、サブピクチャ情報はビットマップ化しているため、外国語文字を含む任意のタイプの文字は、言語ごとにデコーダ内の文字セットを記憶する必要なしに表示することができる。

【0294】サブピクチャ情報の各表示画面が、サブピクチャユニットと呼ばれる。図53にサブピクチャユニット1140がブロック図で示される。サブピクチャユニットは、ビットマップ化されたピクセルデータから成る1つの画面を具備し、このピクセルデータの画面が複数のビデオフレームで表示される。サブピクチャユニット1140は、サブピクチャユニットヘッダ1142を備える。サブピクチャユニットヘッダ1142には、サブピクチャ装置の先頭からのバイト数に相対して、サブピクチャユニットの

サイズおよび表示制御シーケンステーブル1146を記述するフィールドがある。

【0295】ピクセルデータ1144は、ランレンクス圧縮済みビットマップ化情報である。ピクセルは表示された情報のバックグラウンドを形成するバックグラウンドピクセル、図形線描画を含むサブピクチャ内の表示パターンを可能にするパターンピクセル、および2つの異なった属性を持つ図形または文字が表示できるようにする2種類の強調ピクセルとなる場合がある。ビットマップ化情報は、従来のグラフィック作成技法を使用して作成できる。ランレンクス圧縮は、以下の規則に従って実行される。同種の1ピクセルから3ピクセルが続く場合、最初の2ビットにはピクセル数を、続く2ビットにはピクセルデータを入力する。4ビットが1つの単位と見なされる。同種の4から15のピクセルが続く場合は、最初の2ビットにゼロを指定し、続く4ビットにピクセル数、次の2ビットにピクセルデータを入力する。8ビットが1つの単位と見なされる。同種の16から63ピクセルが続く場合、最初の4ビットにゼロを指定し、続く6ビットにピクセル数、次の2ビットにピクセルデータを入力する。12ビットが1つの単位と見なされる。同種の64から255のピクセルが続く場合は、最初の6ビットにゼロを指定し、続く8ビットにピクセル数、接ぎの2ビットにピクセルデータを入力する。16ビットが1つの単位と見なされる。同じピクセルが行末まで続く場合は、最初の14ビットにゼロを指定し、次の2ビットにピクセルデータを記述する。16ビットが1つの単位と考えられる。1行でのピクセルの記述が完了するときにバイト割当てが生じない場合、調整のために4ビットのダミーデータを挿入する。希望に応じて、画像データを表すために他の圧縮機構を使用することも可能である。例えば、JPEGフォーマットやGIFフォーマットを使用して、サブピクチャ画像を表すことができる。

【0296】表示制御シーケンステーブル1146には、制御シーケンスが発生する年代順を記述する複数の表示制御シーケンス1146A、1146B、1146Cなどが含まれる。表示制御シーケンステーブルには、サブピクチャユニット内の情報がどのように表示されるのかに関する情報が含まれる。例えば、1つのサブピクチャを形成するワードは、1つづつまたは経時的にグループとして表示されるか、その色が変化することがある。これは、カラオケ情報を表示するためにサブピクチャを使用する場合に役立つ。

【0297】各表示制御シーケンス(DCSQ)は、サブピクチャユニットが表示されている期間で実行されるピクチャデータを変更するコマンドを酌量する。表示制御シーケンスの第1フィールドは表示開始時間である。このフィールドは、サブピクチャユニットヘッダによりサブピクチャパケット内で記述されるPTSからの相対PTSのあるDCSQに含まれる表示制御コマンドの実行開始時間を記述

する。記述された実行開始時間の後の第1ビデオフレームから、表示制御緒は、DCSQ表示制御開始時間に従って開始される。各DCSQの第2フィールドは、後続の表示制御シーケンスのアドレスである。このフィールドは、第1サブピクチャユニットからの相対バイト数により後続のDCSQの開始アドレスを記述する。後続のDCSQが存在しない場合には、このフィールドは、第1サブピクチャユニットからの相対バイト数によりこのDCSQの開始アドレスを記述する。それ以降、DCSQには1つ以上の表示制御コマンドが格納される。これらのコマンドにより、ピクセルデータの属性および表示を制御、変更できるようになる。コマンドには、サブピクチャのオン/オフ状態に関係なく、強制的にサブピクチャ装置の表示を開始するコマンドが含まれる。例えば、ユーザがサブピクチャをオフにするためにメニューを使用する場合、このコマンドを設定してユーザの設定値を無視することができる。別のコマンドは、更新されたサブピクチャ装置の表示を開始する。このコマンドは、各DCSQで一度表示されなければならない。2種類の強調ピクセル、パターンピクセル、およびバックグラウンドピクセルを含む前記の4種類のピクセルの色とコントラストを設定するためのコマンドがある。サブピクチャを構成するピクセルデータの矩形領域表示および位置を設定するコマンドが存在する。このコマンドは、上部および下部のX座標とY座標の設定を可能にする。また、表示に使用されるピクセルデータのヘッドアドレスを設定するために使用されるコマンドもある。このコマンドを使用すると、上部フィールドと下部フィールドの両方のヘッドアドレスを入力できる。拡大されたフィールドでは、装置のヘッドからの相対バイト数による上部フィールドと下部フィールドのピクセルデータの各ヘッドアドレスが使用される。第1ピクセルデータは、行の第1ピクセルを表す。このコマンドは、少なくとも第1DCSQ、DCSQ0で使用されなければならない。

【0298】ピクセルデータの色とコントラストを変更するためのDCSQの中に表示されるコマンドがある。データ検索情報のハイライト情報が使用される場合には、このコマンドは使用してはならない。ピクセルデータの色とコントラストを変更するためのコマンドは、表示中にピクセルデータの色とコントラストを制御するためのピクセル制御データを含む。ピクセル制御データに記述される制御の内容は、指定された開始時間の後にビデオフレームごとに実行され、新規ピクセル制御データが検出されるか、新規サブピクチャユニットが検出されるまで続行する。ピクセル制御データは、同じ変更が加えられる線を指定するために使用される線制御情報を含む。また、ピクセル制御情報は、変更が加えられる線上の位置を指定するために使用することもできる。線制御コマンドにより、開始線番号、変更ポイント数、および変更終了線番号の変更が可能になる。ピクセル制御情報は、変

更開始ピクセル番号、および強調ピクセル1と2、パターンピクセル、およびバックグラウンドピクセルの新しい色とコントラストを含む。また、表示制御シーケンスの最後を示すコマンドもある。各DCSQは、このコマンドで終了しなければならない。

【0299】映画の場合のようなビデオフレームの期間で、使用できる多くの異なったサブピクチャユニットが存在する。図54に図示されるように、サブピクチャユニットは、サブピクチャパック1150に分割される。サブピクチャパックには、前記のパックヘッダ972、MPEGシステム要件に準拠するパックヘッダ、後続のデータがサブピクチャデータであることを示すサブストリームID、およびサブピクチャデータ1158それ自体がある。

【0300】サブピクチャユニット1140のようなサブピクチャユニットは、図55にブロック図で示されるようにサブピクチャパック150に分割される。最後のサブピクチャパック1050Gには、その長さを2048バイトにする埋込みが含まれる。

【0301】図56では、連続サブピクチャユニットが表示される方法を説明する。時間が新しいサブピクチャパケットに記述されるPTSの時間に等しくなる場合、現在表示中(1160)のサブピクチャユニットがクリアされ(1162)、次のサブピクチャユニットの表示制御シーケンスによって指定される時間が発生すると、そのサブピクチャが表示される(1164)。1160と1162の間に、サブピクチャユニットnとサブピクチャユニットn+1の間の境界線がある。この境界線は、サブピクチャユニットn+1のパケットヘッダに記述されるPTSで発生する。サブユニットn+1の実際の表示時間は、サブピクチャn+1のDCSQに説明される。

【0302】サブピクチャ情報に使用されるピクセルデータの作成は、表示が希望される情報のビットマップを作成するためにコンピュータを使用して実行できる。ビットマップ化情報のプロセスは、技術で周知である。

【0303】図40-56は、オーディオ情報、ビデオ情報、およびサブピクチャ情報をエンコードするために使用されるデータ構造を強調する。ただし、図40-56でのデータ構造の開示により、MPEGエンコードの従来の技術の当業者は、特にMPEGシステム説明ISO/IEC 13818-1の力を借りて、エンコードされたオーディオ、ビデオ、およびサブピクチャをデータ構造にフォーマットできる。同様に、データを記憶する構造についての知識があれば、ビデオ情報、オーディオ情報、およびサブピクチャ情報を作成するためのデータ構造の復号化は、従来の技術の当業者により実行できる。

【0304】本発明の一部を成す例示的なビデオデコーダは、光ディスクに記憶されたデータを読み取る光ディスク読取り装置を具備する。読み取られた情報は、情報を復号化するために、従来の解析技法を使用して解析される。ビデオ、オーディオ、サブピクチャ、およびPBI

パックのすべてを復号化しなければならない。ビデオは、オーディオパックが従来のデコーダを使用して復号化されるように、市販されているMPEGデコーダを使用して復号化することができる。サブピクチャ情報は、サブピクチャユニットをサブピクチャパックから、それらがエンコードされたのと逆転して構築することで復号化される。特別に構築されたハードウェアまたはソフトウェアコーディングを使用してプログラミングされた汎用マイクロプロセッサは、サブピクチャ情報を復号化するために使用できる。図45にブロック図で示される再生情報パックには、データ検索情報996が含まれる。データ検索情報は、プログラミングされたマイクロプロセッサを使用して処理され、オーディオデコーダ、ビデオデコーダ、およびサブピクチャデコーダとは異なる。メニュー機能もプログラミングされたマイクロプロセッサによって実行されるので、データ検索情報の中に含まれるハイライト情報1022および他のすべてのメニュー機能は、サブピクチャデコーダによってではなく、プログラミングされたマイクロプロセッサによって実行される。本発明の動作、データ構造、エンコードプロセスと復号化プロセスに関する追加情報は、両方とも本明細書に参考により取り入れられる日本国特許明細書7-8298、および7-85693に記載される。

【0305】本発明のエンコードシステムを使用すると、ピクチャ内で品質を制御できる。この制御は、最終ビデオ製品の作成者がエンコードシステムの最終製品を密接に制御できるようになるため重要である。したがって、デジタルビデオエンコードによりアナログビデオ内に存在しない復号化人為構造が生じても、最終デジタル製品はきわめて良好だろう。前記項のどれかの教示は、他の項に適用できる。

【0306】本発明の決定、計算および処置は、コンピュータ技術の当業者に明らかとなるように、本発明の教示に従ってプログラミングされた従来の汎用デジタルコンピュータを使用して、実現できる。適切なソフトウェアコーディングは、ソフトウェア技術の当業者に明らかとなるように、本開示の教示に基づいて技能のあるプログラマにより容易に作成できる。

【0307】本発明は、当業者に容易に明らかになるように、アプリケーションに特殊な集積回路を作成したり、従来の構成部品回路の適切なネットワークの相互接続によっても、実現される。

【0308】本発明は、本発明のプロセスを実行するためにコンピュータをプログラミングする場合に使用できる指示を格納する記憶装置媒体であるコンピュータプログラム製品を含む。記憶装置媒体は、フロッピーディスク、光ディスク、CD-ROM、および磁気光学ディスク、ROM、RAM、EPROM、EEPROM、磁気カードや光カード、あるいは電子指示を記憶するために適当な任意のタイプの媒体を含む任意のタイプのディスクを含むが、それらに限

定されない。本発明は、さらに、電子指示またはデータを記憶するのに適当な前記媒体のどれかで記憶される、本発明により出力されるエンコードでデータを含む記憶装置媒体であるコンピュータプログラム製品も含む。前記教示という観点から、本発明の数多くの改良および変化が可能であることは明らかである。したがって、付属請求項の範囲内であれば、本発明は、本明細書に明確に指示がある場合を除き実施できる。

【0309】

【発明の効果】以上説明したようにこの発明によれば、編集したときに画質を劣化させないようにすることができる。即ち、エンコードビデオのセクションを除去するため、エンコードされたビットストリーム内のエンコードビデオの始点及び終点を決定する。エンコードデータの始点及び終点アドレスは、各ピクチャーで消費されたビット、各ピクチャーのシーケンスヘッダーのビット数、各ピクチャーグループ (GOP) ヘッダーのビット数、及び全スタッフビットを加算することにより計算される。除去されるビデオの代わりに挿入されるビデオをエンコードするため、無関係な信号成分のデコードを防ぐ処理が実行される。この処理は編集ポイント前の最終ピクチャーを判断し、編集ポイント後のフレームに関する参照フレームとしてその最終ピクチャーを使用することで行われる。更に、その最初の量子化値を用いてエンコードされた最初のビデオの短い期間が、置き換えられるビデオの終わりに含まれる。又は、初期にエンコードされた参照ピクチャーが使用されないように、エンコーダは動作する。しかしこれは、ピクチャーの品質を維持するために余分のビットが必要となる。なぜなら、置き換えビデオの初期の双方向フレームは2つの参照フレームの代わりに1つの参照フレームのみを使用するからである。

【図面の簡単な説明】

【図1A】エンコーダシステムアーキテクチャのブロック図。

【図1B】エンコードされるべきオーディオおよびビデオを得るための装置、ならびにエンコードシステムにより作成されるエンコードされたオーディオ映像データを書き込み、伝送するための装置を示す図。

【図2】図1のシステムの一般的な動作を示すフローチャート。

【図3】セットアップ状態で実行されるプロセスを示すフローチャート。

【図4】圧縮状態で実行されるプロセスを示す図。

【図5】編集状態で実行されるプロセスを示す図。

【図6】開始状態、完了状態、および終了状態のために実行されるプロセスを示す図。

【図7】マルチパスビデオエンコーダを使用しながら所定数のオーディオトラックをエンコードするのに必要とされるオーディオエンコーダの数を決定する方法、なら

びにオーディオエンコードおよびマルチパスビデオエンコードを同時に実行するプロセスを示すフローチャート。

【図8A】シーケンスレーヤのログファイルフォーマットを示す図。

【図8B】同じくシーケンスレーヤのログファイルフォーマットを示す図。

【図8C】同じくシーケンスレーヤのログファイルフォーマットを示す図。

【図8D】同じくシーケンスレーヤのログファイルフォーマットを示す図。

【図9A】ピクチャレイヤのログファイルフォーマットを示す図。

【図9B】同じくピクチャレイヤのログファイルフォーマットを示す図。

【図9C】同じくピクチャレイヤのログファイルフォーマットを示す図。

【図10】マクロブロックレーヤのログファイルのフルフォーマット、およびマクロブロックレーヤのログファイルのショートフォーマットを示す図。

【図11】ログファイル内の情報を利用するためのプロセスを示す図。

【図12】異なる期間に対する画質の手動調整中に実行される一般プロセスを示す図。

【図13】様々な期間で画質を変更するためのパラメータを入力するために使用されるユーザインタフェースを示す図。

【図14】所望の画質に対応するビット数を計算するための一般的な手段を示す図。

【図15】最大または最小のユーザ選択優先順位が設定されるセクションを処理するために使用されるプロセスを示す図。

【図16A】最大または最小の優先順位ではない希望の画質に対応するビット数を計算するためのフローチャートを示す図。

【図16B】同じく最大または最小の優先順位ではない希望の画質に対応するビット数を計算するためのフローチャートを示す図。

【図17】画質を変更した結果、アンダフローが生じる場合に確認するために使用されるプロセスのフローチャートを示す図。

【図18】バッファアンダフローが、編集セグメント内で加えられた変更のために、編集セグメントの後のフレームで発生するかどうかを判断するためのフローチャートを示す図。

【図19】ユーザが選択した品質によってエンコードされたビデオのビットレートがどのように変更されるのかについての例、および同じくユーザが選択した品質によってエンコードされたビデオのビットレートがどのように変更されるのかについての例を示す図。

【図20】1つのフレーム内のさまざまな領域の画質の手動調整の間に実行される一般的なプロセスを示す図。

【図21】ビデオの1フレーム内にユーザが選択した優先度が設定されたさまざまな領域の例を示す図。

【図22】選択された優先順位に使用される元の量子化の端数小数部に対するユーザ選択優先順位のグラフを示す図。

【図23】ユーザ選択優先順位の結果生じるビット数と、フレームの元のエンコードの結果生じるビット数の間の差異が許容できるかどうか、あるいは訂正手順を実行しなければならないかどうかを判断するためのプロセスを示す図。

【図24A】設定されたビットが多すぎると判断されるために、ビット数を低減する目的でマクロブロックの量子化レベルが増分される場合に、フレームを訂正するためのプロセスを示す図。

【図24B】同じく設定されたビットが多すぎると判断されるために、ビット数を低減する目的でマクロブロックの量子化レベルが増分される場合に、フレームを訂正するためのプロセスを示す図。

【図24C】同じく設定されたビットが多すぎると判断されるために、ビット数を低減する目的でマクロブロックの量子化レベルが増分される場合に、フレームを訂正するためのプロセスを示す図。

【図25A】ビット数を増加するために1つのフレーム内のマクロブロックの量子化レベルを引き下げることによってフレーム内の少なすぎるビットを訂正するためのプロセスを示す図。

【図25B】同じくビット数を増加するために1つのフレーム内のマクロブロックの量子化レベルを引き下げることによってフレーム内の少なすぎるビットを訂正するためのプロセスを示す図。

【図25C】同じくビット数を増加するために1つのフレーム内のマクロブロックの量子化レベルを引き下げることによってフレーム内の少なすぎるビットを訂正するためのプロセスを示す図。

【図26】 n 個のマクロブロックの集合のそれぞれで使用するビットの予想数をメモリからロードしてから、 n 個のマクロブロックの集合ごとにビットの結果数をカウントし、ビットの予想数および作成数に基づいた概算の精度を計算し、フレームの列のターゲットサイズに近く留まるために、次の n 個のマクロブロックの予め割り当てられた量子化レベルを調整する一般的なプロセスを示しているフローチャート。

【図27】(A)は、固定割当てされた量子化レベルを示しているフレームのピクチャを示す図、(B)はマクロブロックの各集合に1スライスが含まれ、マクロブロックの第1集合が再エンコードされた後で、第1訂正係数が計算され、第1訂正係数がマクロブロックの第2集合に加算された場合の、図27(A)のフレームのピク

チャを示す図、(C)はマクロブロックの第2集合が再エンコードされた後で、第2訂正係数が計算され、第2訂正係数がマクロブロックの第3集合に追加された、図27(B)のフレームのピクチャを示す図、(D)はマクロブロックの第1集合に2つのスライスが含まれ、マクロブロックの第1集合が再エンコードされてから、第1訂正係数が計算され、第1訂正係数がマクロブロックの第2集合に加算された場合の図27(A)のフレームのピクチャを示す図。

【図28】その量子化レベルが変更される2つの異なった領域を含む1フレームを示す図と、その量子化レベルが変更された2つの異なった領域を表すために使用されるビデオストリーム内でのビット数の変更を示している図28(A)のフレームの2つの考えられるエンコードのピクチャを示す図。

【図29】再エンコードされた列の総長が列の元の長さ に等しい、一連の画像を表すために使用されるビット数の変更を示しているビデオストリームを示す図。

【図30】訂正係数指数が、どのようにして概算率を使用して比例ルックアップテーブルから求められるのかを説明するフローチャートを示す図。

【図31】訂正係数が、訂正係数指数を使用して訂正係数テーブルからどのようにして選択されるのかを説明するフローチャートを示す図。

【図32】訂正係数が、概算割合およびマクロブロックの再エンコードされていない残りの集合数を使用してどのようにして計算されるのかを説明するフローチャートを示す図。

【図33】(A)はエンコードされたフォーマットを取るオリジナルビデオのセクションを示す図、(B)は元のエンコードされたビデオに置き換えられなければならないエンコードされていないフォーマットにおけるビデオの新たなセクションを示す図、(C)はオリジナルビデオのエンコードされていないセクションがその両端にあるビデオのエンコードされていない新たなセクションを示す図、(D)は元のエンコードされたビデオに置き換えられなければならないエンコードされたフォーマットにおける図33(C)のビデオを示す図、(E)は置き換えられたエンコードされたビデオがその中に含まれる、元のエンコードされたビデオを示す図。

【図34】減少されたデコード構造を有するビデオとなるエンコードビデオデータストリームにビデオのセクションを置き換えるプロセスを示す図。

【図35】(A)は立ち下がり端でのオリジナルビデオのエンコードされていないセクションを有するビデオのエンコードされていない新しいセクションを示す図、

(B)は図33(A)のブロック図にある最初にエンコードされたビデオに置き換えられるエンコードされたビデオのセクションを示す図、(C)は図35(B)の置き換えられたエンコードされたビデオを含む図33

(A) の最初にエンコードされたビデオを示す図。

【図 3 6】 クローズグループのピクチャモードを使用して置き換えられるビデオをエンコードすることによって、エンコードされたビデオビットストリームを編集する場合に、デコード構造を減少させるためのプロセスを説明する図。

【図 3 7】 (A) は事前エンコードプロセスでビットレートモデルに対する量子化レベルを作成するために、それに割り当てられる2つの量子化レベルが設定される未処理ビデオのフレームを示す図、(B) は回転パターンで割り当てられる4つの量子化レベルのあるフレームを示す図、(C) はブロックフォーメーションで分散される4つの量子化レベルのあるフレームを示す図。

【図 3 8】 ビデオ信号を得て、それらの信号を事前エンコードし、事前エンコードフェーズで特定された量子化レベルに対するビットレートを決定するプロセスを説明するフローチャート。

【図 3 9】 2つの記録済みフレーム、およびビットレートに対するそれらの対応する量子化レベルを示す図。

【図 4 0】 最終的には光ディスクに格納されるエンコードされたデータの最終フォーマットの概略を示す図。

【図 4 1】 図 4 0 のディスク情報ファイル 908 を説明する図。

【図 4 2】 図 4 0 にブロック図で示されるデータファイルのデータファイル管理情報の内容を示す図。

【図 4 3】 図 4 0 のデータファイルに記憶されるデータの詳細を示す図。

【図 4 4】 その中でそのデータがスタッフィングパケットを必要としないデータパックを示す図とパディングのためにパケットを必要とするデータパックを示す図。

【図 4 5】 再生情報パックを示す図。

【図 4 6】 図 4 5 のデータ検索情報 996 の中に含まれる情報を示す図。

【図 4 7】 図 4 6 の一般情報を示す図。

【図 4 8】 図 4 6 にブロック図で示される同期再生情報を示す図。

【図 4 9】 ビデオパックを示す図。

【図 5 0】 ピクチャのグループ (GOP) とビデオパックの列の間の関係を説明する図。

【図 5 1】 MPEGオーディオエンコード仕様に従ってエンコードされたオーディオパックを示す図と、AC-3 またはリニア PCM オーディオエンコードに従って構築されたオーディオパックを示す図。

【図 5 2】 エンコードされたオーディオストリームとオーディオパックの間の関係を示す図。

【図 5 3】 サブピクチャユニットの構造を示す図。

【図 5 4】 サブピクチャパックの構造を示す図。

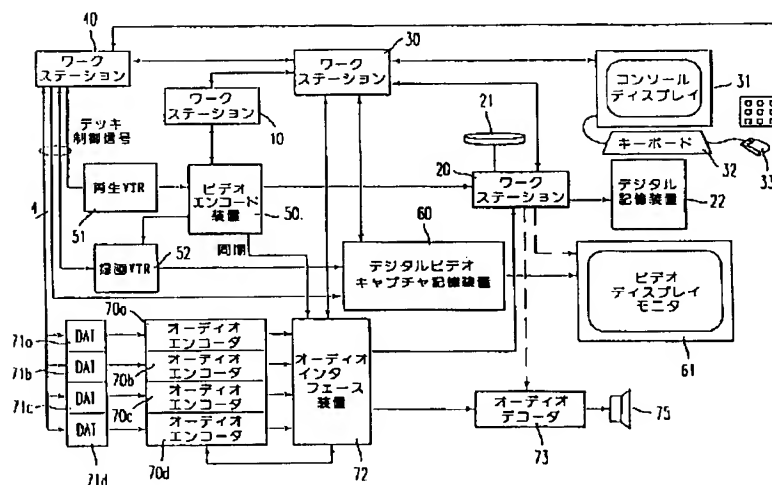
【図 5 5】 サブピクチャパックに関連したサブピクチャ装置の構造を説明する図。

【図 5 6】 表示サブピクチャユニット n とサブピクチャユニット n+1 の間の変化を示す図。

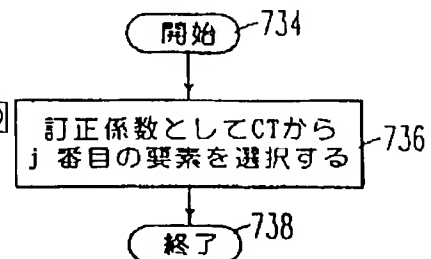
【符号の説明】

10、20、30、40…ワークステーション、21…ハードディスク、22…デジタル記憶装置、31…コンソールディスプレイ、50…ビデオエンコード装置、51…再生 VTR、52…録画 VTR、60…デジタルキャプチャ記憶装置、70a~70d…オーディオエンコーダ、72…オーディオインタフェース装置、73…オーディオデコーダ。

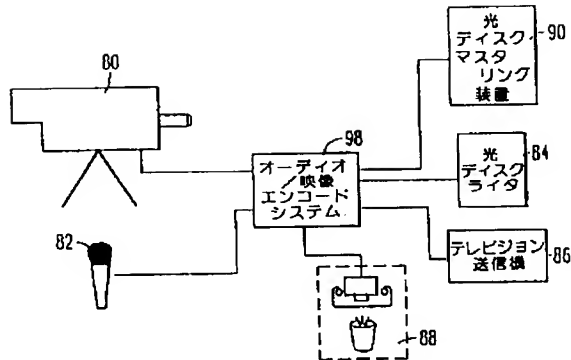
【図 1 A】



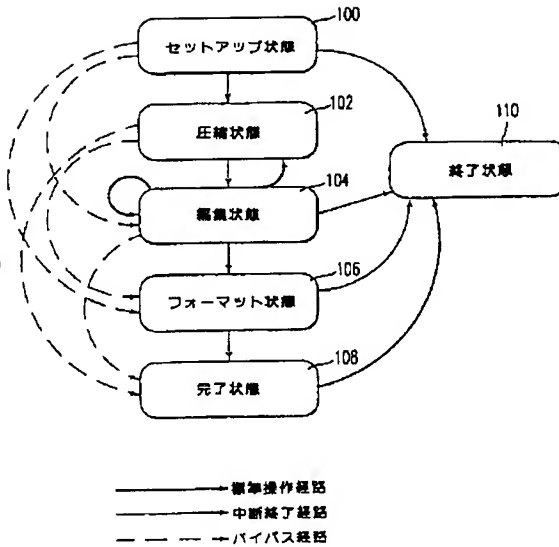
【図 3 1】



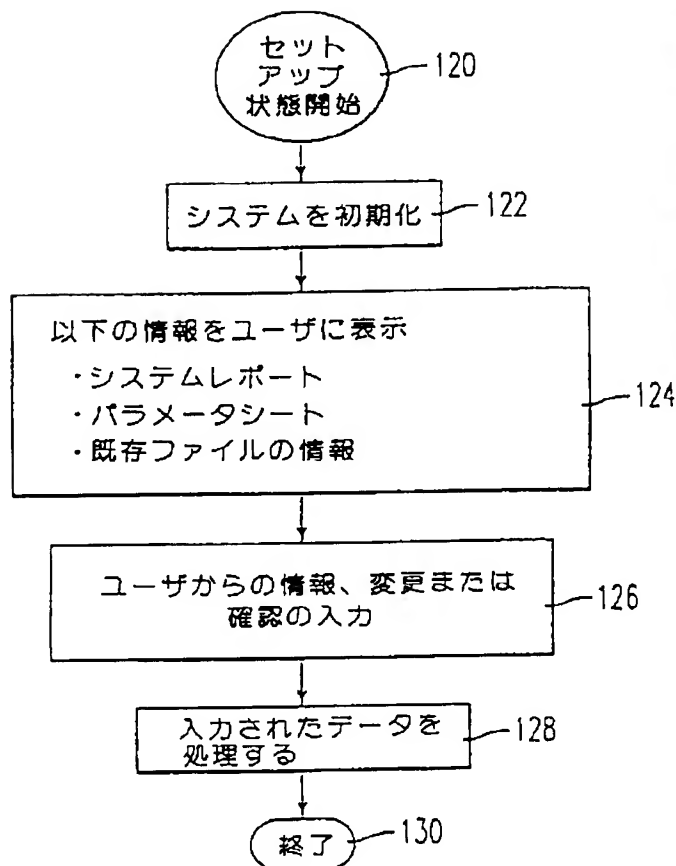
【図1B】



【図2】



【図3】

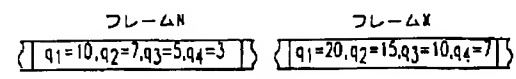


【図13】

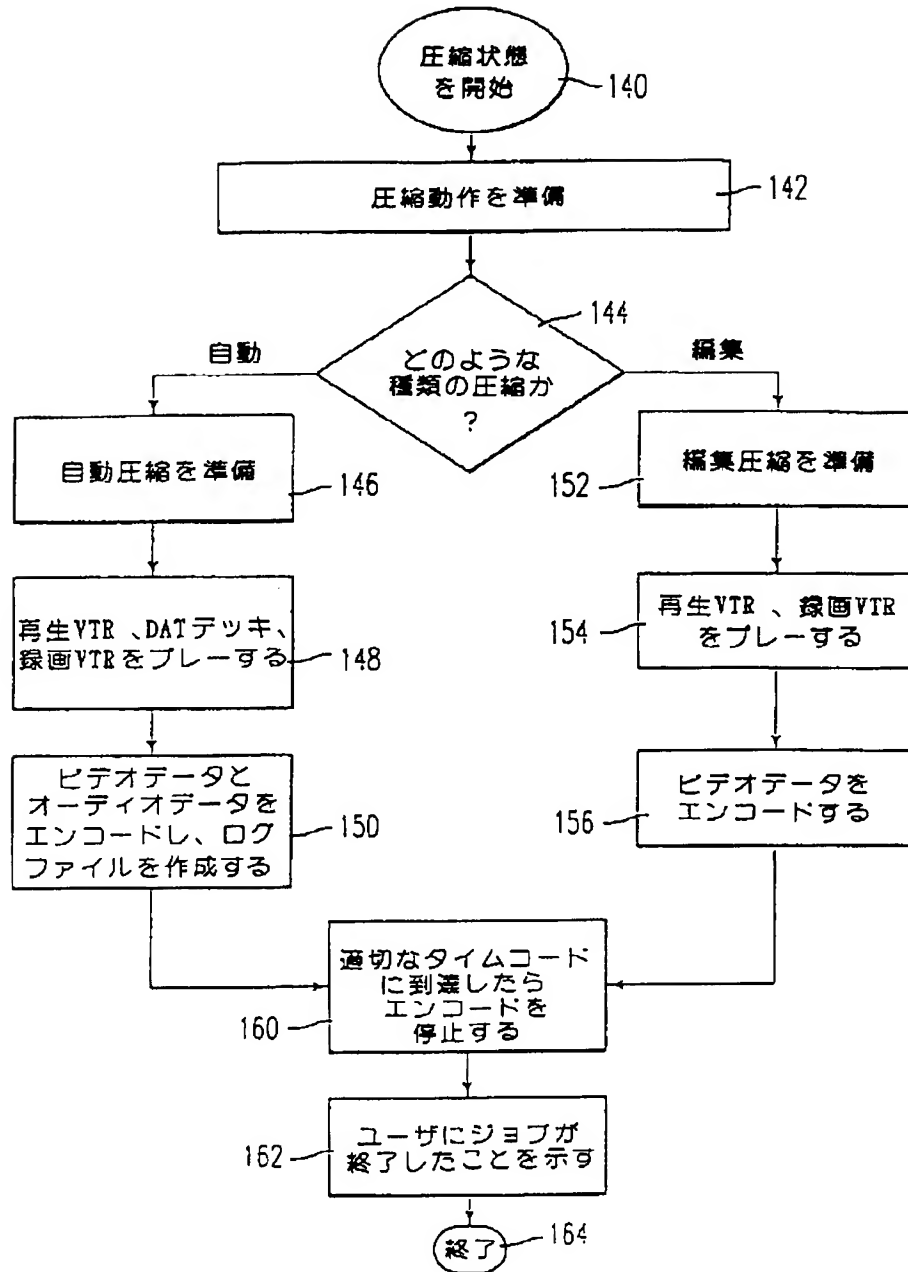
イン	アウト	優先順位
00:00:50:00	00:01:00:00	P
00:01:00:01	00:02:00:00	-1
00:02:07:00	00:03:20:00	-1
00:04:00:00	00:04:50:00	3

イン	00:00:00:00	アウト	00:00:00:00	優先順位	P	=
最小ビットレート	5,678	バックグラウンド優先順位	P	=		

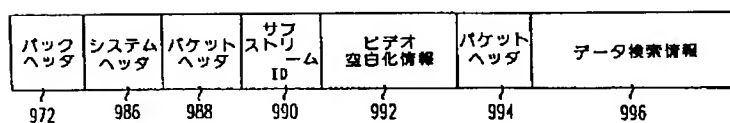
【図39】



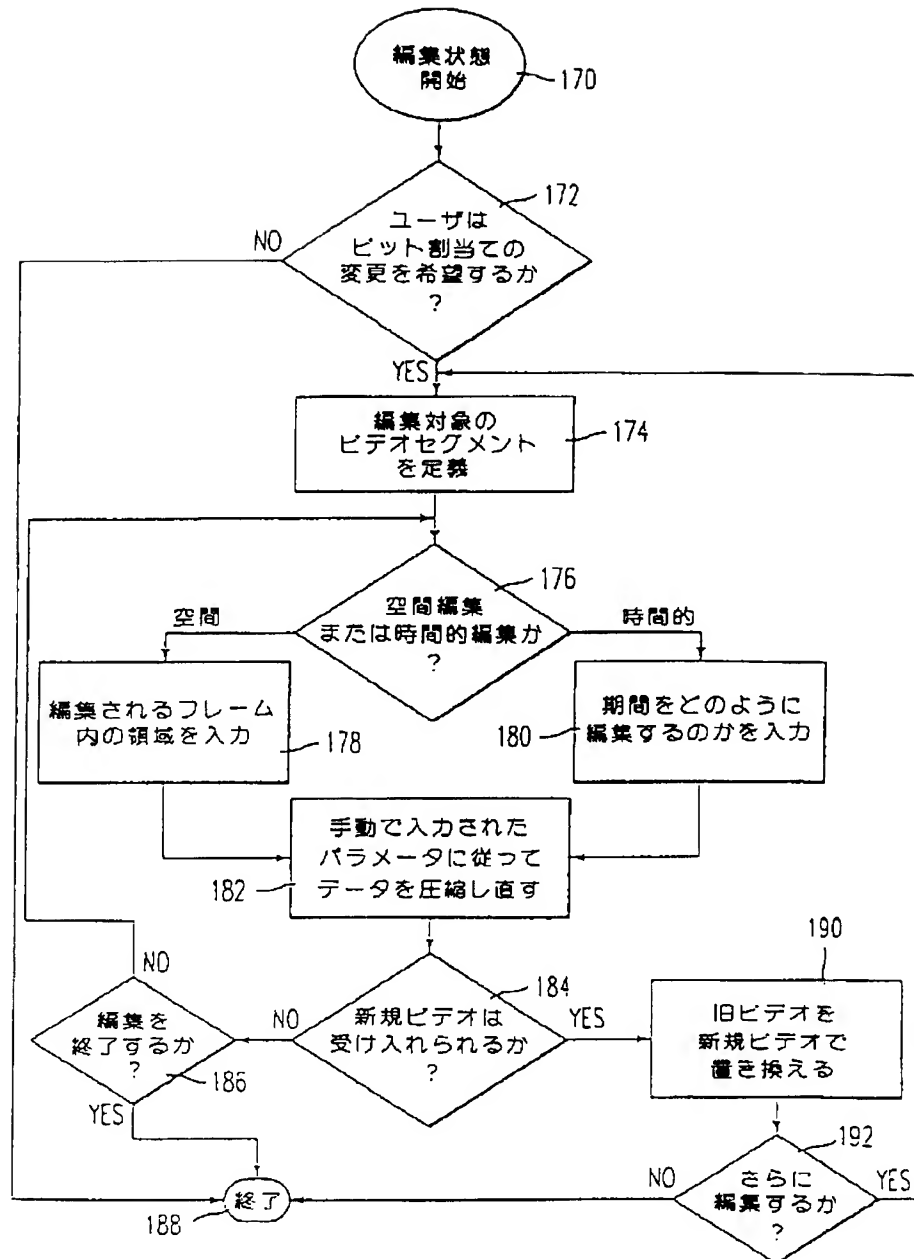
【図4】



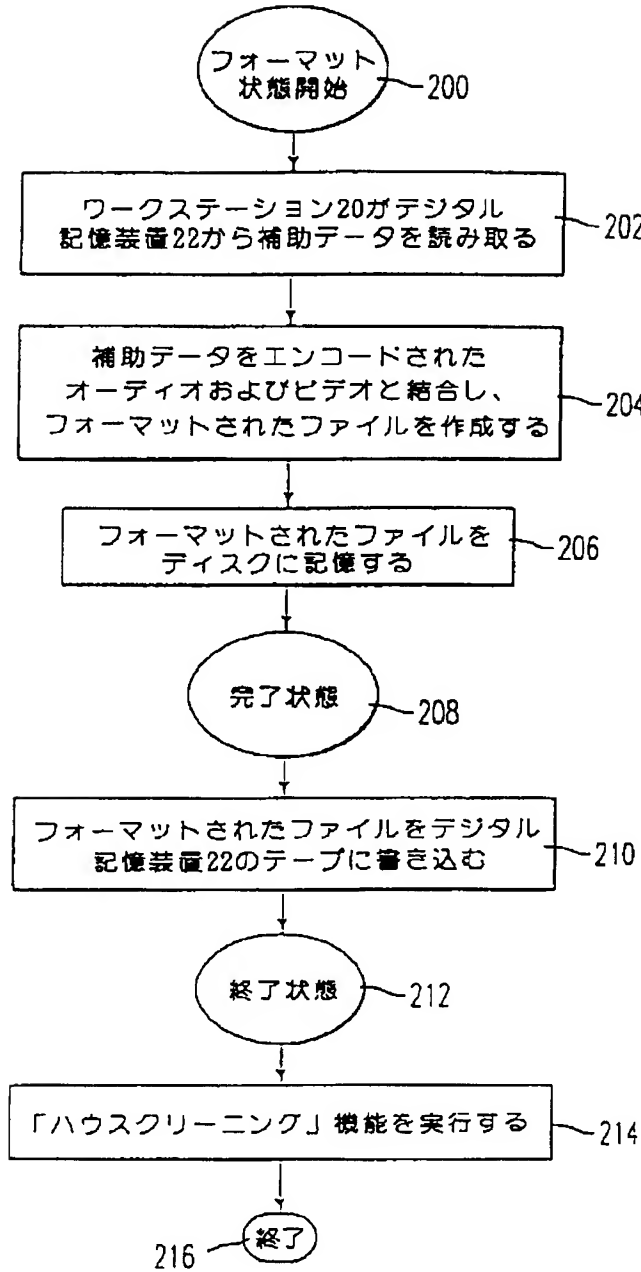
【図45】



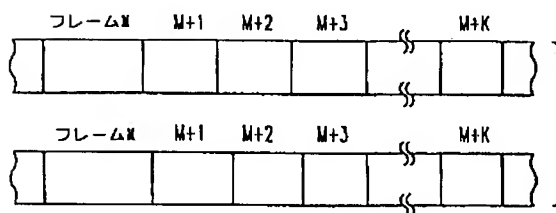
【図5】



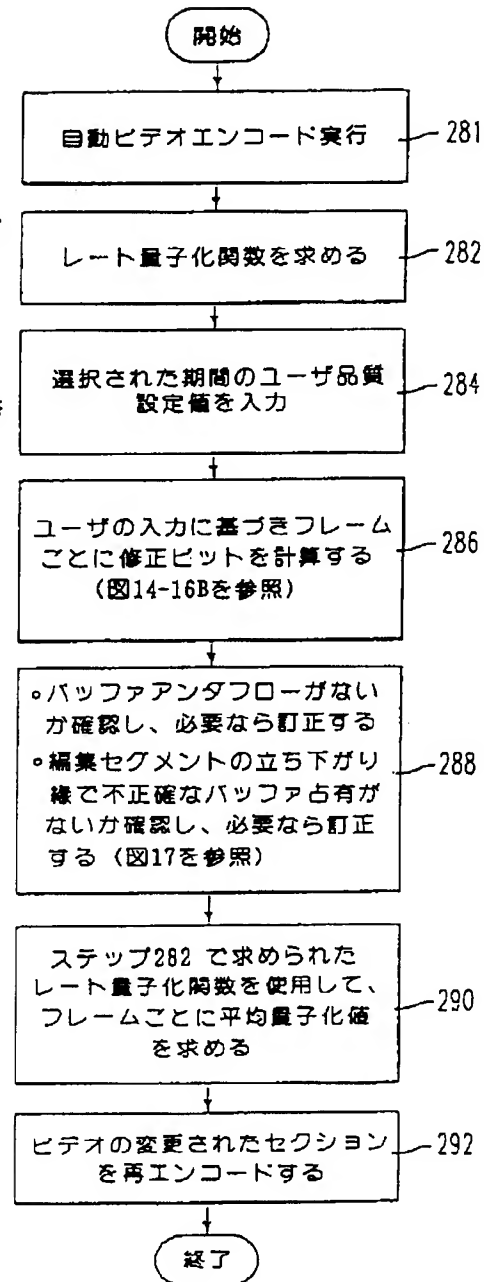
【図6】



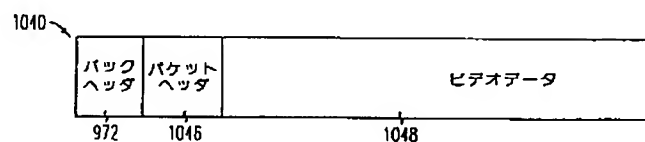
【図29】



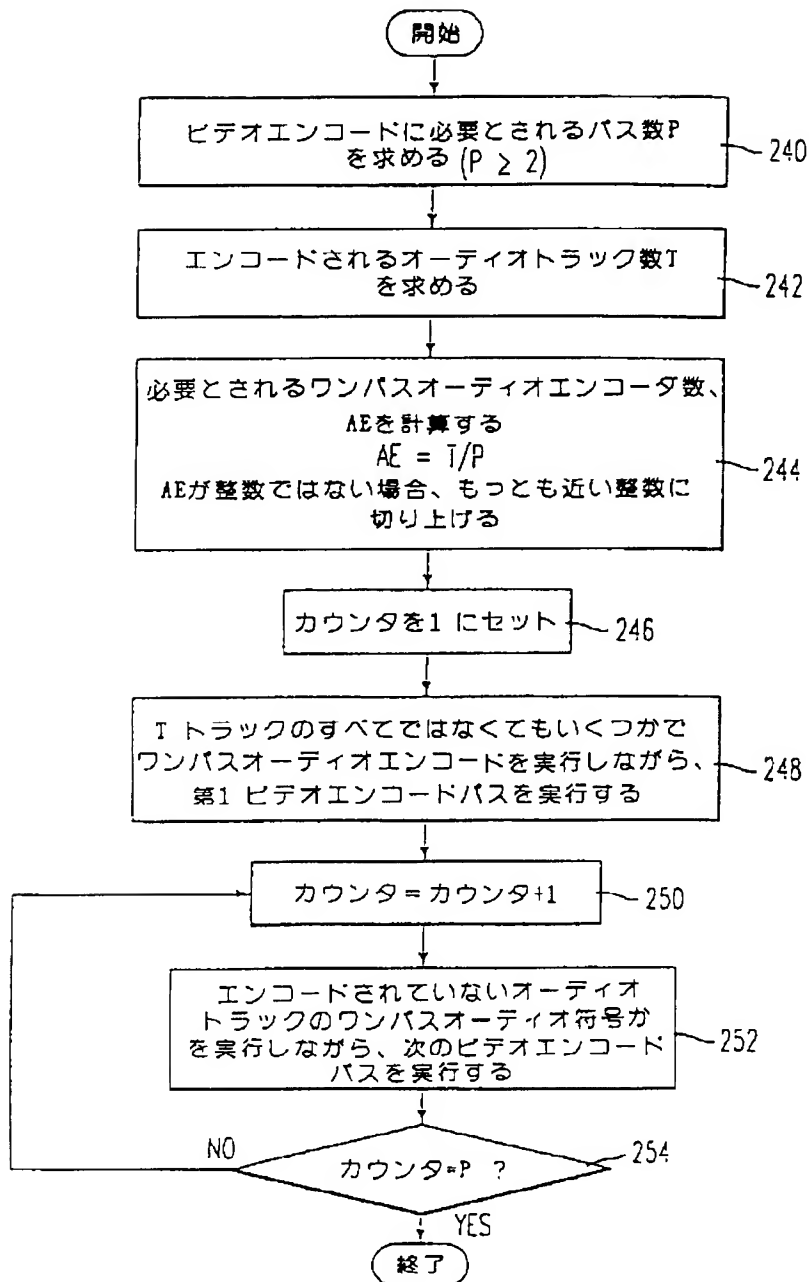
【図12】



【図49】



【図7】

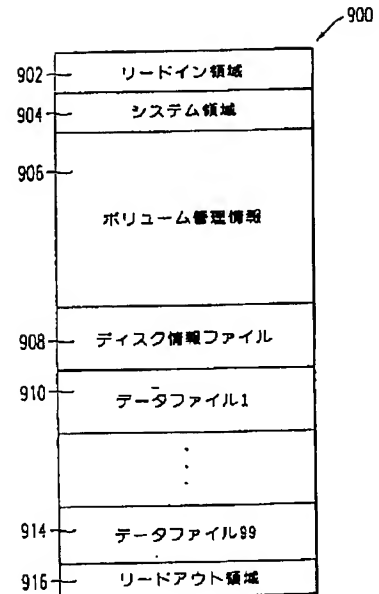


【図8A】

	ビット	ビット フィールド	内容
version	8	[31..20]	パラメータフォーマットバージョン
begin_lc	24	[23..0]	開始タイムコード
end_lc	24	[31..0]	終了タイムコード
ntsc_pal	1	[7..7]	ntsc : 0, pal : 1
low_delay	1	[6..6]	低ディレー : 1, 他 : 0
enable_3_2	1	[5..5]	3-2プルダウン使用 : 1, 他 : 0
MPEG1_2	1	[4..4]	MPEG1 : 0, MPEG2 : 1
MCmode	1	[3..3]	固定 : 0, 順応 : 1
Mmax	2	[2..1]	最大サイズM
cbr_vbr	1	[0..0]	CBR : 0, VBR : 1
MD_hsize	6	[31..26]	画像のMB幅
MD_vsize	6	[25..20]	画像のMB高さ
initial_delay	15	[19..6]	イニシャルバッファフルネス
word_boundary2	5	[5..0]	ワードバンタリーの反転
bit_rate	24	[31..8]	平均ビットレート
n	2	[7..6]	Mが可変ならば、Mは0
n	6	[5..0]	Nが可変ならば、Nは0
VBV_size	24	[31..8]	ビットのVBVバッファサイズ
word_boundary4	0	[7..0]	ワードバンタリーの保存
number_of_coded_picture	32	[31..0]	総コード化画像の数

FIG. 8Bへ

【図40】



【図8D】

FIG. 8Cから

frame_memory_channel	2	[3..2]	0 : CH1, 1 : CH2, 2 : スルー
timecode_check_enable	1	[1..1]	1 : イネーブル
word_boundary76	1	[0..0]	ワードバンタリーの保存
frame_memory_file_num	5	[31..27]	フレームメモリファイル数
initial_pulldown_th	10	[26..9]	初期プルダウンスレッシュホルド
word_boundary77	9	[8..0]	ワードバンタリーの保存
scene_change_threshold	32	[31..0]	シーンチェンジ順応
pulldown_phase_threshold_file	8x256		プルダウンフェーズのファイル名
MD_ctrl_file	8x256		MBコントロールのファイル名
pon_scan_file	8x256		パンスキャンデータのファイル名
VBR_parameter_file	8x256		VBRパラメータのファイル名
M_sequence_file	8x256		Mシーケンスのファイル名

【図8B】

FIG. 8Aから

intro_q1[64]	8x64	[7..0]	イントラ q-matrix 1
intro_q2[64]	8x64	[7..0]	イントラ q-matrix 2
non_intro_q1[64]	8x64	[7..0]	非イントラ q-matrix 1
non_intro_q2[64]	8x64	[7..0]	非イントラ q-matrix 2 (ジグザグスキャン順)
total_bits[2]	32x2	[31..0]	ビットストリームの総数 最上位32ビット ([0]) 最下位32ビット ([1])
max_bit_rate	24	[31..0]	VBRモードに対する最大ビットレート
word_boundary5	0	[7..0]	ワードバンダリーの保存
Ki	0	[31..24]	レートコントロール定数
Kp	0	[23..16]	レートコントロール定数
Kb	0	[15..8]	レートコントロール定数

FIG. 8Cへ

【図9C】

FIG. 9Bから

stuffing_size	20	[19..0]	画像のゼロスタッフィングサイズ
bits_picture_header	12	[31..20]	ピクチャーヘッダビット
bits_MB_header	20	[19..0]	スライスヘッダ、MBヘッダ DC及びEOBビット
word_boundary10	2	[31..30]	ワードバンダリーの保存
mean_activity	30	[29..0]	メニューアクティビティ
mc_error	32	[31..0]	MCエラー
coding_error	32	[31..0]	コーディングエラー
S1_I	32	[31..0]	レートコントロール状態
S2_I	32	[31..0]	レートコントロール状態
S1_P	32	[31..0]	レートコントロール状態
S2_P	32	[31..0]	レートコントロール状態
S1_B	32	[31..0]	レートコントロール状態
S2_B	32	[31..0]	レートコントロール状態
Qsum	32	[31..0]	Qの合計
reserved(4)	32	[31..0]	

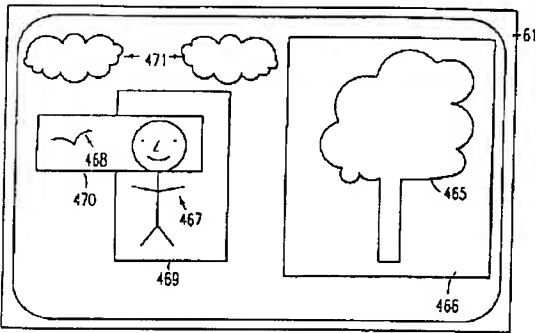
【図8C】

FIG. 8Bから

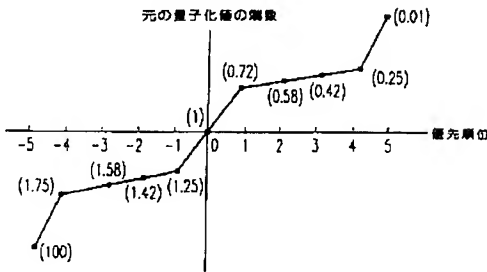
	ビット	ビット フィールド	内容
rc_type	8	[7..0]	レートコントロールタイプ
Ki2	8	[31..24]	レートコントロール定数
Kp2	8	[23..16]	レートコントロール定数
Kb2	8	[15..8]	レートコントロール定数
rc_type2	8	[7..0]	レートコントロールタイプ2
display_hsize	14	[31..18]	表示水平サイズ
display_vsize	14	[17..4]	表示垂直サイズ
word_boundary75	4	[3..0]	ワードバウンダリーの保存
run_mode	2	[31..30]	ランニングモード 0 : ロープコントロール 1 : タイムコードコントロール 2 : タイムスタート及びループコントロール
loops	10	[29..12]	ループコントロール中の特定ループ
pulldown_load_mode	2	[11..10]	プルダウンフェーズを獲得方法 0 : 自動, 1 : ログから 2 : データファイルから
scene_change_adp_enable	1	[9..9]	1 : シーンチェンジ順応オン
sequence_header_mode	1	[8..8]	1 : 各GOPヘッダの前
pon_scan_enable	1	[7..7]	1 : イネーブル
HB_parameter_mode	1	[6..6]	1 : フル, 0 : ショート
buffer_check_enable	2	[5..4]	0 : ディスエーブル 1-3 : イネーブル

FIG. 8Dへ

【図21】



【図22】



【図46】

996

一般情報	1000
アングルアドレス情報	1020
ハイライト情報	1022
周囲再生情報	1024
PBI バックアドレス情報	1036

992

【図9A】

	ビット	フィールド	内容
time-code	24	[31..8]	タイムコード
picture_coding-type	2	[7..6]	画像コーディングタイプ
top_field_first	1	[5..5]	トップフィールドファースト
repeat_first_field	1	[4..4]	リピートフィールド
display_repeat_field	1	[3..3]	表示画の繰り返し数
load_intra_qmat	1	[2..2]	ロードイントラ量子化マトリックス
load_non_intra_qmat	1	[1..1]	ロード非イントラ量子化マトリックス
pan_scan	1	[0..0]	使用 picture_display_extension()
input_order	32	[31..0]	コード化画像の入力順序
prefil_coeff	4	[31..28]	プリフィルタ係数インデックス
nr_coeff	2	[27..26]	NR係数インデックス
nc_rm_type	2	[25..24]	nc_rm_type インデックス
rlc_zero_type	2	[23..22]	rlc_zero_type インデックス
vbv_occupancy	21	[21..1]	MPEG2DISにおけるVBVオキュパンシオン
HLmode_type	1	[0..0]	インターレース：0、非インターレース：1
forward_ae	23	[31..9]	フォワードレファレンスのMEエラー
mean_qscales	8	[8..1]	平均αスケール×4
processing_type	1	[0..0]	4：2：2→4：2：0フィルタタイプ 0：フィールド／1：フレーム

FIG. 9B

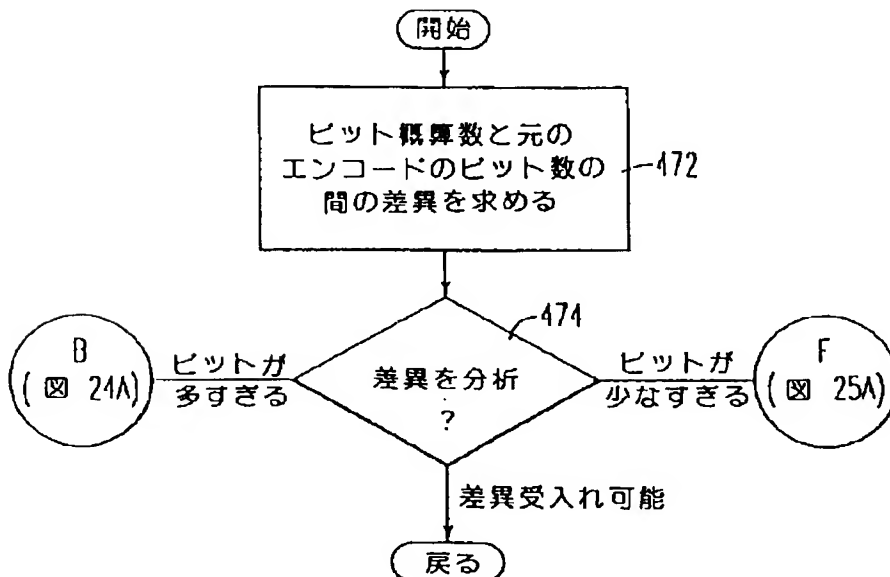
【図9B】

FIG. 9Aから

backward_ue	23	[31..9]	バックワードリファレンスのNEエラー
intra_q_matrix_type	1	[8..8]	イントラ量子化マトリックスタイプ
non_intra_q_matrix_type	1	[7..7]	非イントラ量子化マトリックスタイプ
intra_dc_precision	2	[6..5]	イントラDC精度
q_scale_type	1	[4..4]	qスケールタイプ
concealment_MVs	1	[3..3]	コンシールメントMV
intra_vlc_format	1	[2..2]	イントラVLCフォーマット
alternate_scan	1	[1..1]	交互スキャン
progressive_frame	1	[0..0]	プログレッシブフレーム
allocated_bit	21	[31..11]	画像のアロケートビット
temporal_reference	10	[10..1]	テンポラルリファレンス
scene_change	1	[0..0]	1: シーンチェンジ
generated_bit_picture	21	[31..11]	画像用生成ビット
word_boundary6	11	[10..0]	ワードバンダリーの保存
word_boundary6	11	[10..0]	ワードバンダリーの保存
			シーケンスヘッダビット
GOP_number	20	[19..0]	GOP番号
bits_GOP_header	12	[31..20]	画像がある場合の、GOPヘッダビット

FIG. 9Cへ

【図23】



【図10】

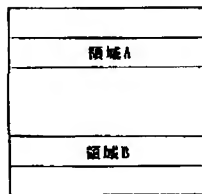
		ビット		フィールド	内容
(A)	{	MBパラメータ	ビット		
		word_boundary0	2	[31..30]	ワードバンダリーの保存
		MB_type	5	[29..25]	マクロブロックタイプ
		MC_type	2	[24..23]	MCタイプ
		DCT_type	1	[22..22]	DCTタイプ
		header_length	8	[21..14]	MBヘッダ、DCビット
		code_length	14	[13..0]	RLCビット
		word_boundary2	1	[31..31]	ワードバンダリーの保存
		MB_act	15	[30..16]	MBアクティビティ
		mc_error_MB	16	[15..0]	MCエラー
		q_scale	5	[31..27]	量子化スケールインデックス
		reserved	27	[26..0]	ワードバンダリーの保存

		ビット		フィールド	内容
(B)	{	MBパラメータ	ビット		
		MB_type	5	[31..27]	マクロタイプ
		header_length	8	[26..19]	MBヘッダ、DCビット
		code_length	14	[18..5]	RLCビット
		q_scale	5	[4..0]	量子化スケールインデックス

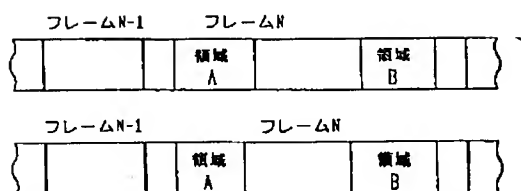
【図28】

(A)

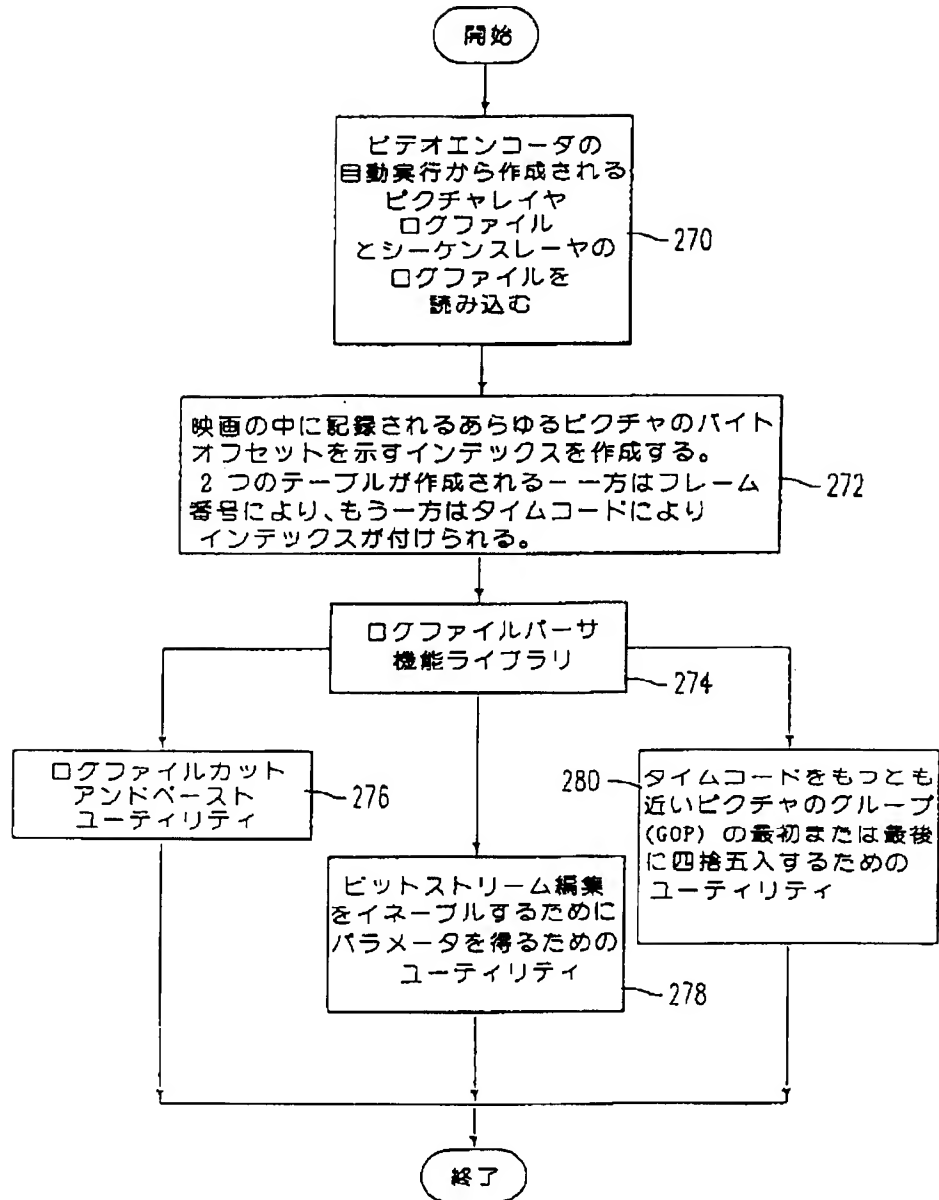
フレームNの表示



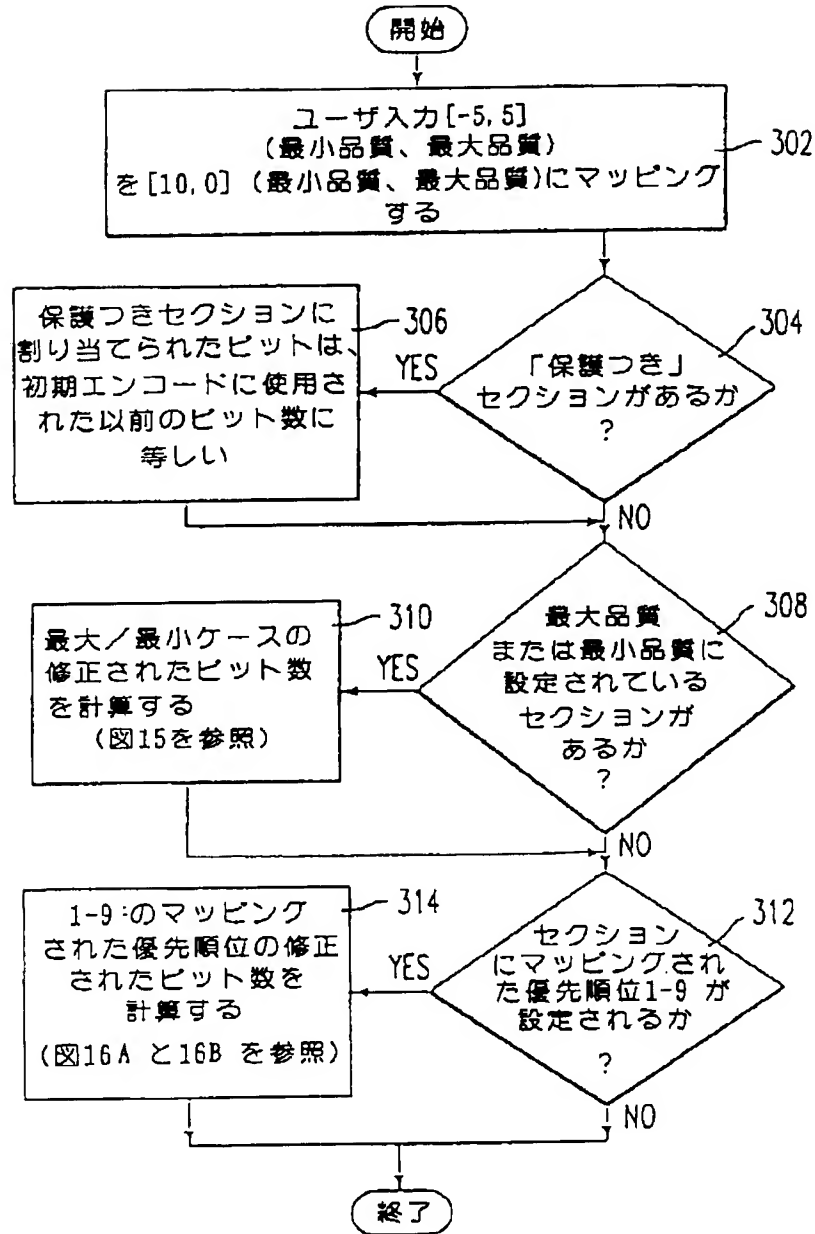
(B)



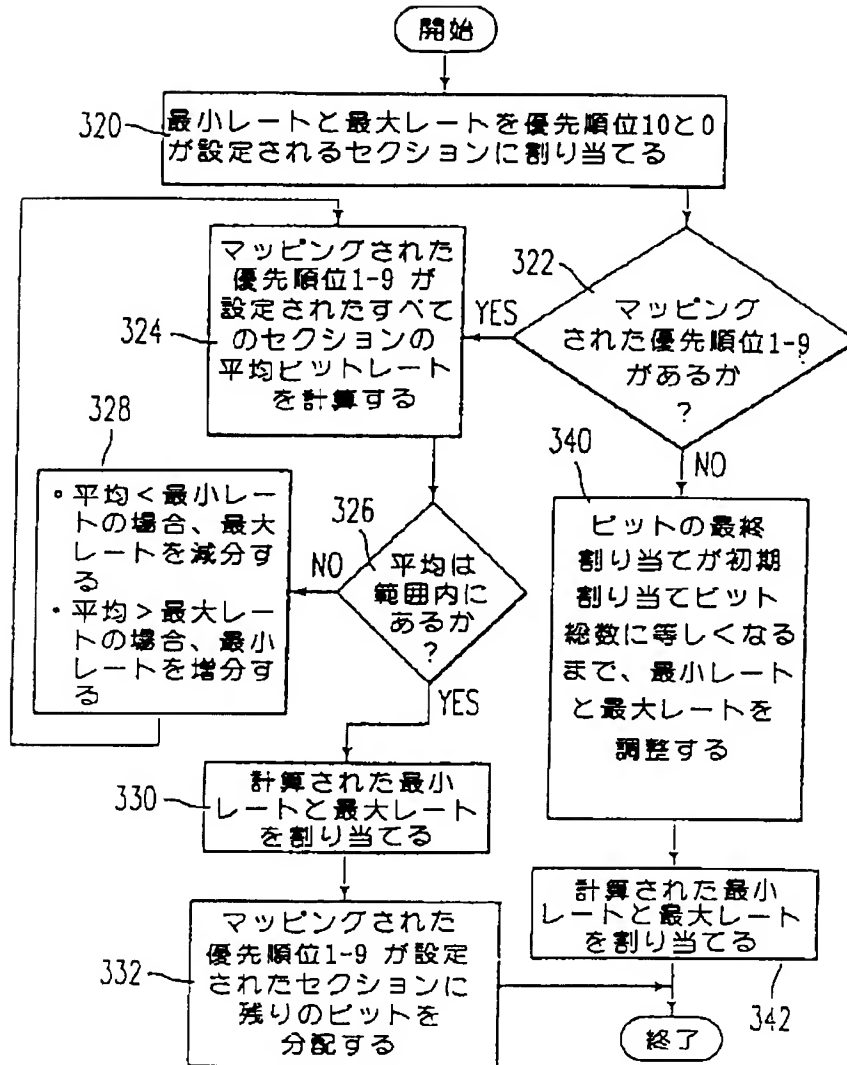
【図11】



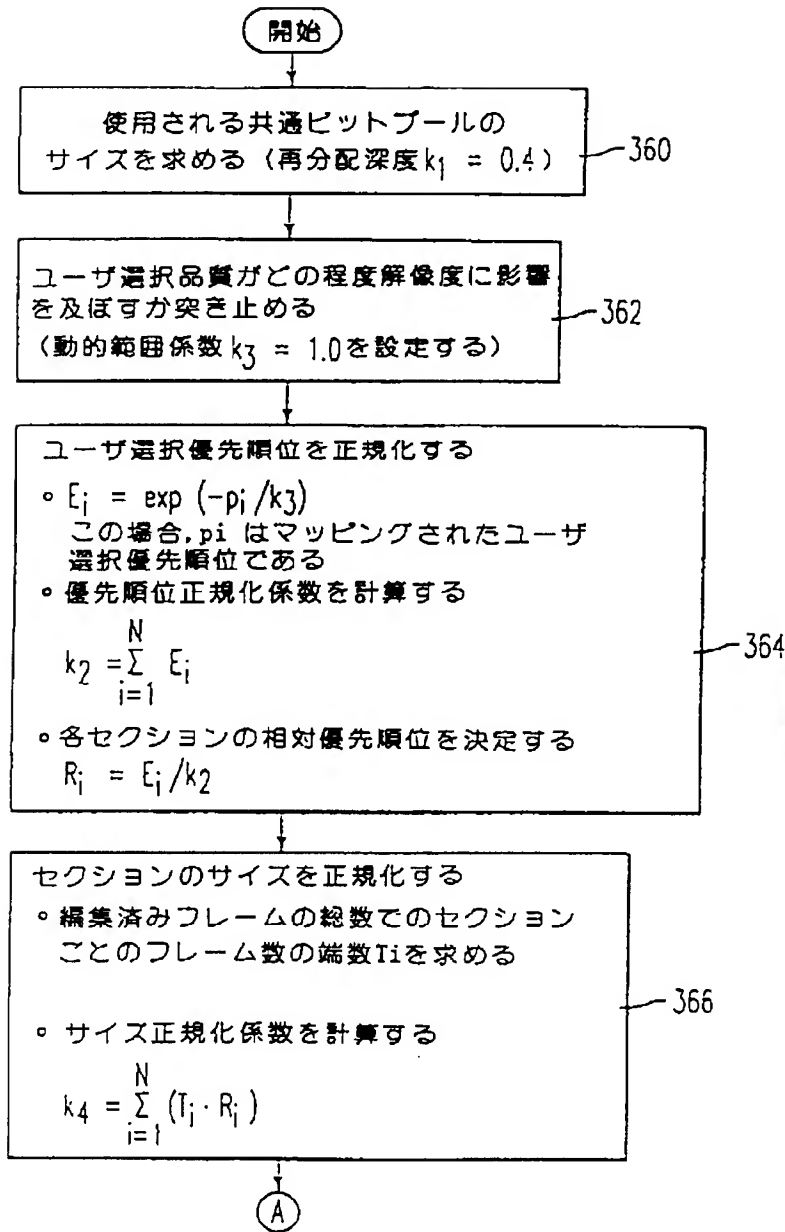
【図14】



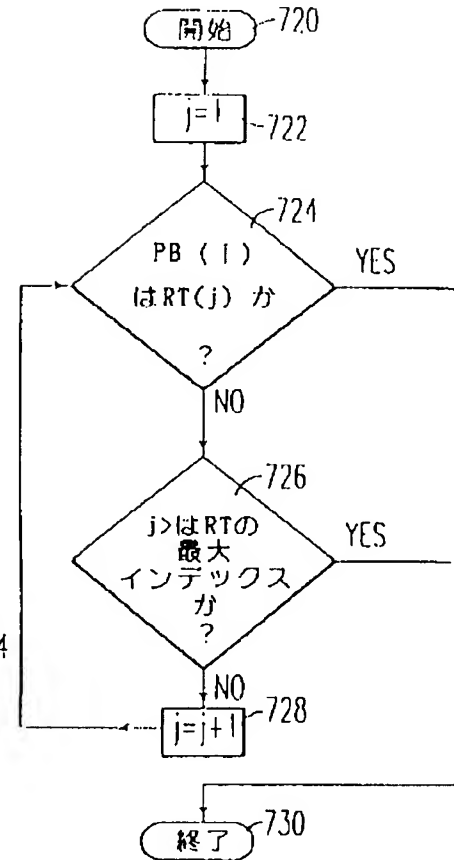
【図15】



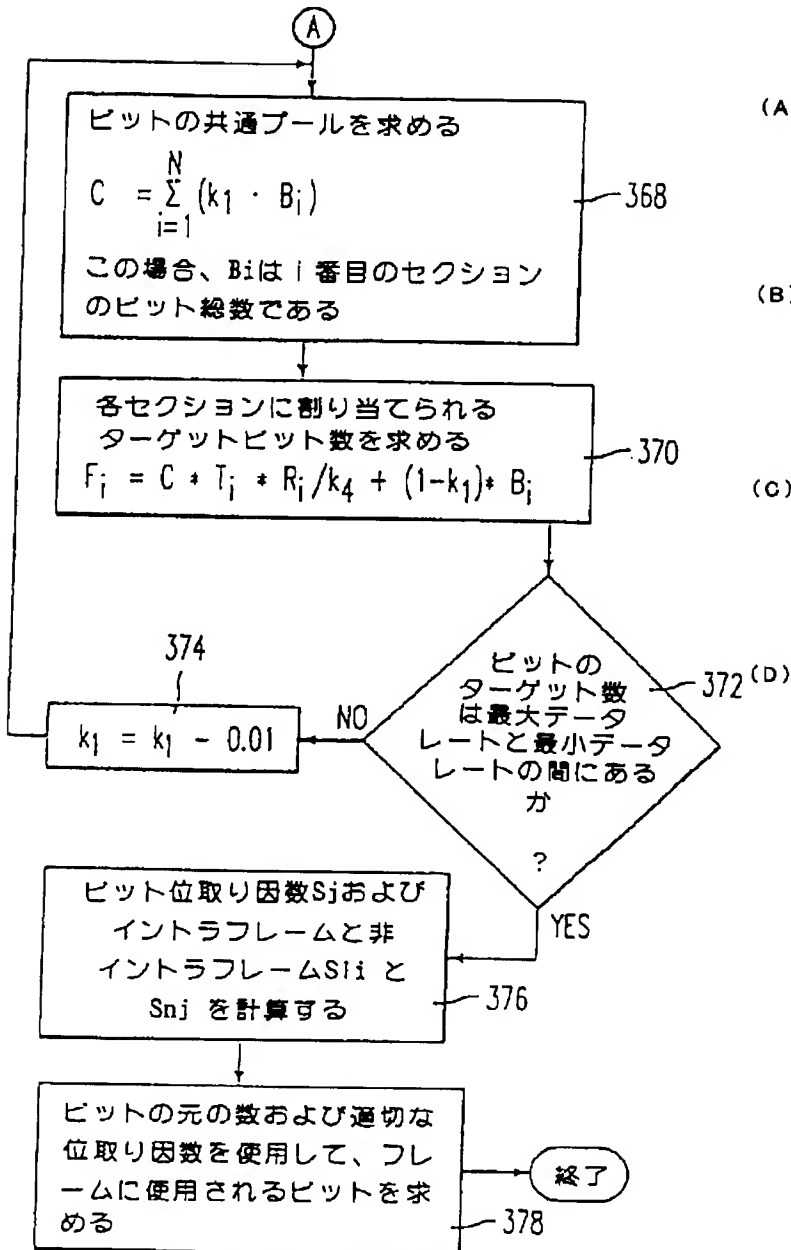
【図16A】



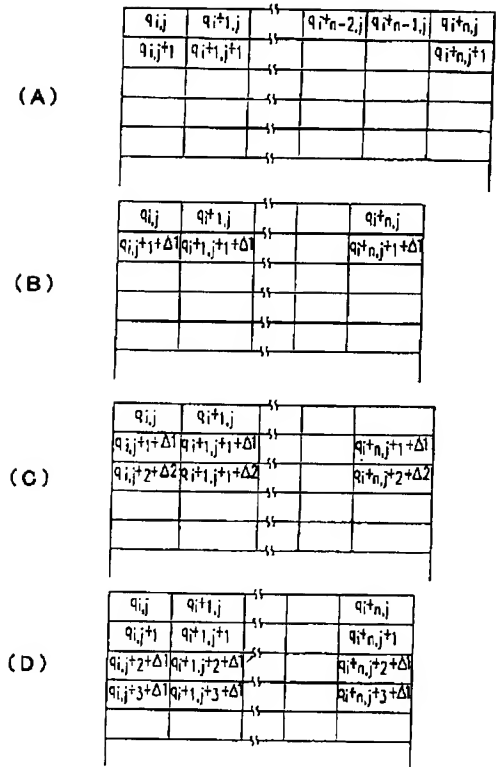
【図30】



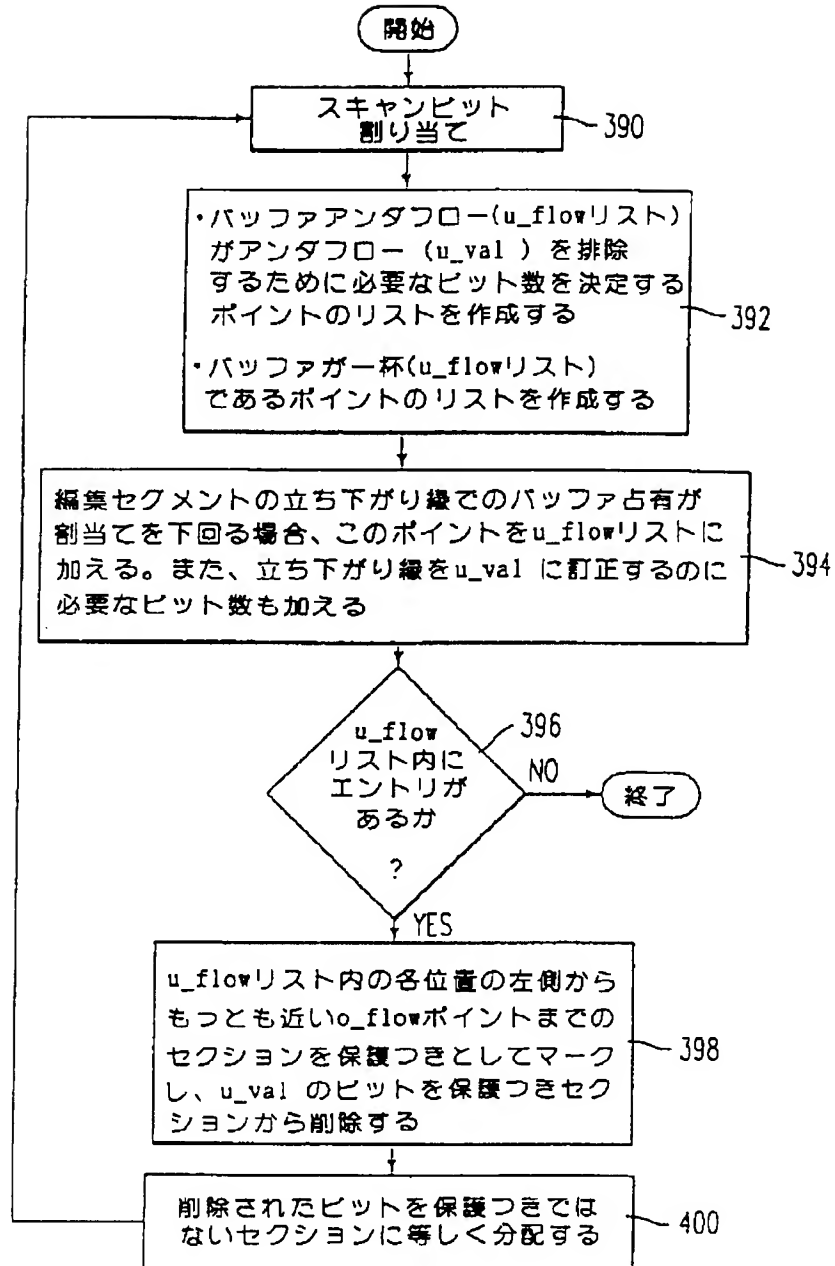
【図16B】



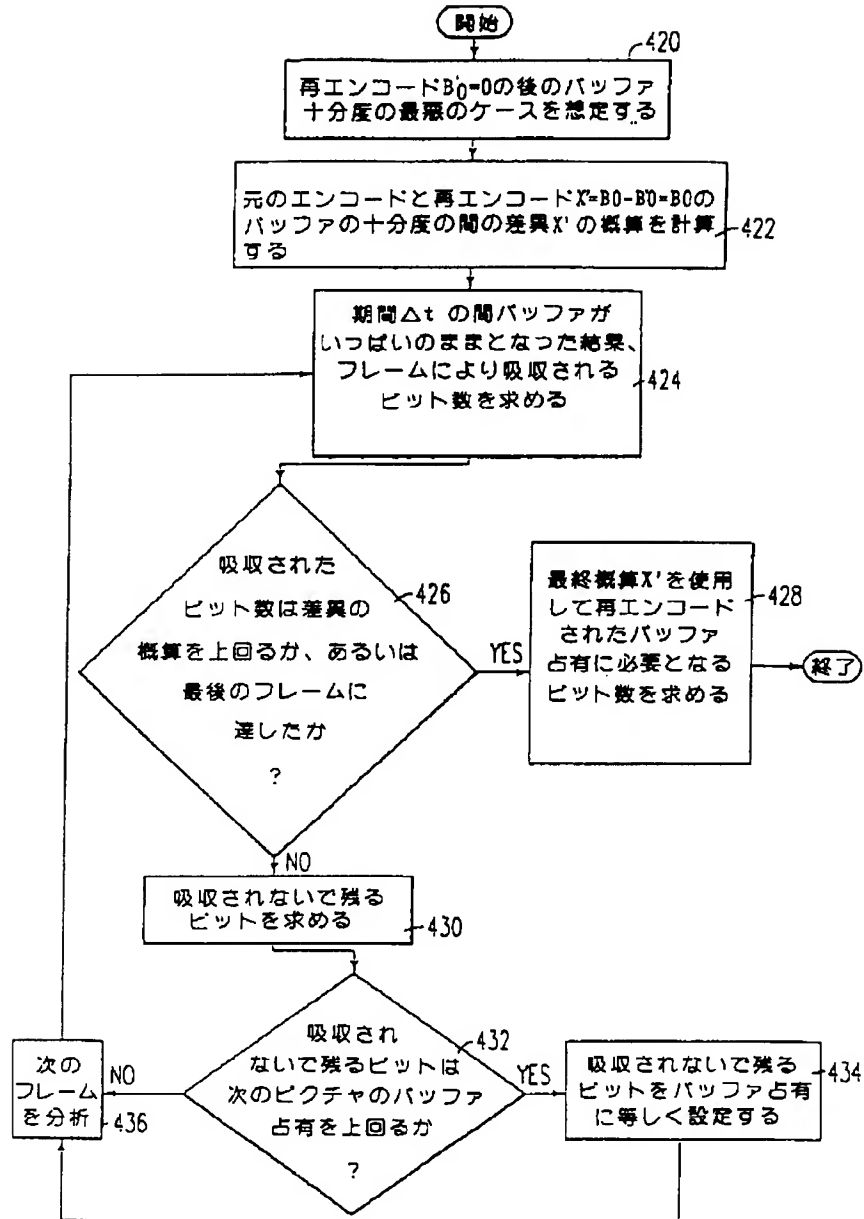
【図27】



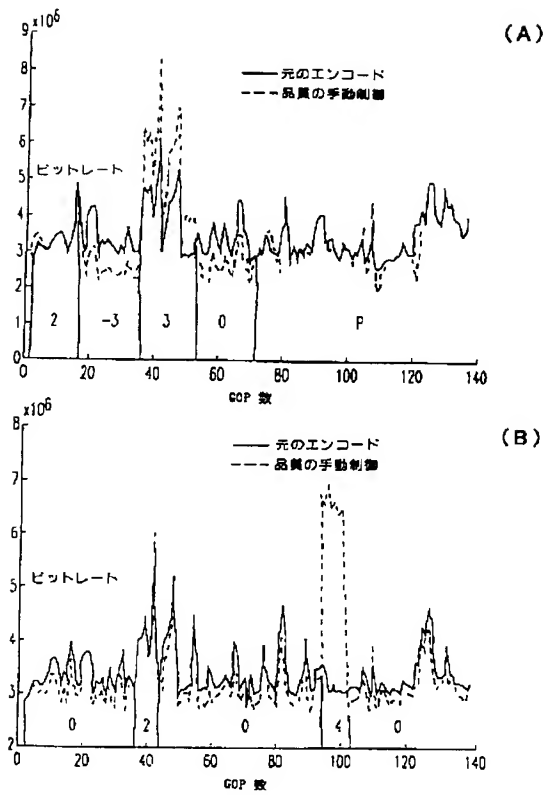
【図17】



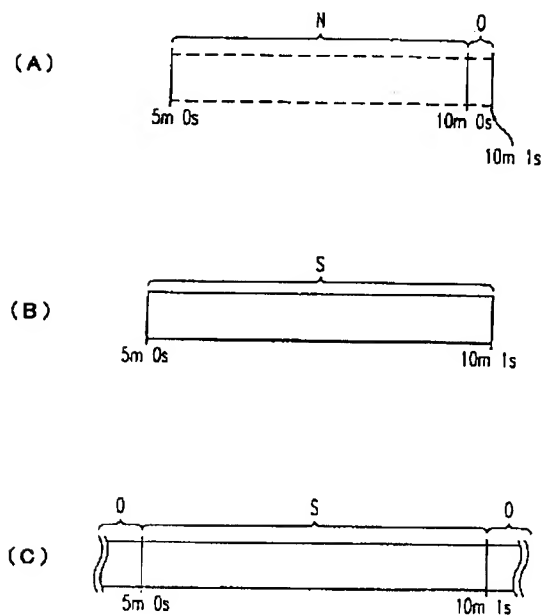
【図18】



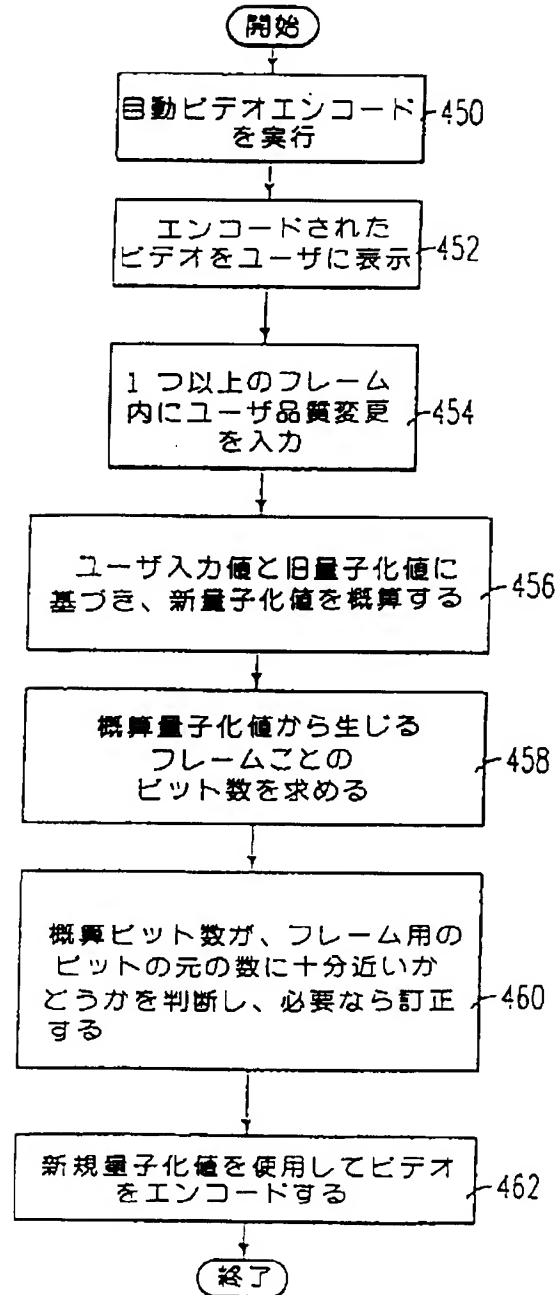
【図19】



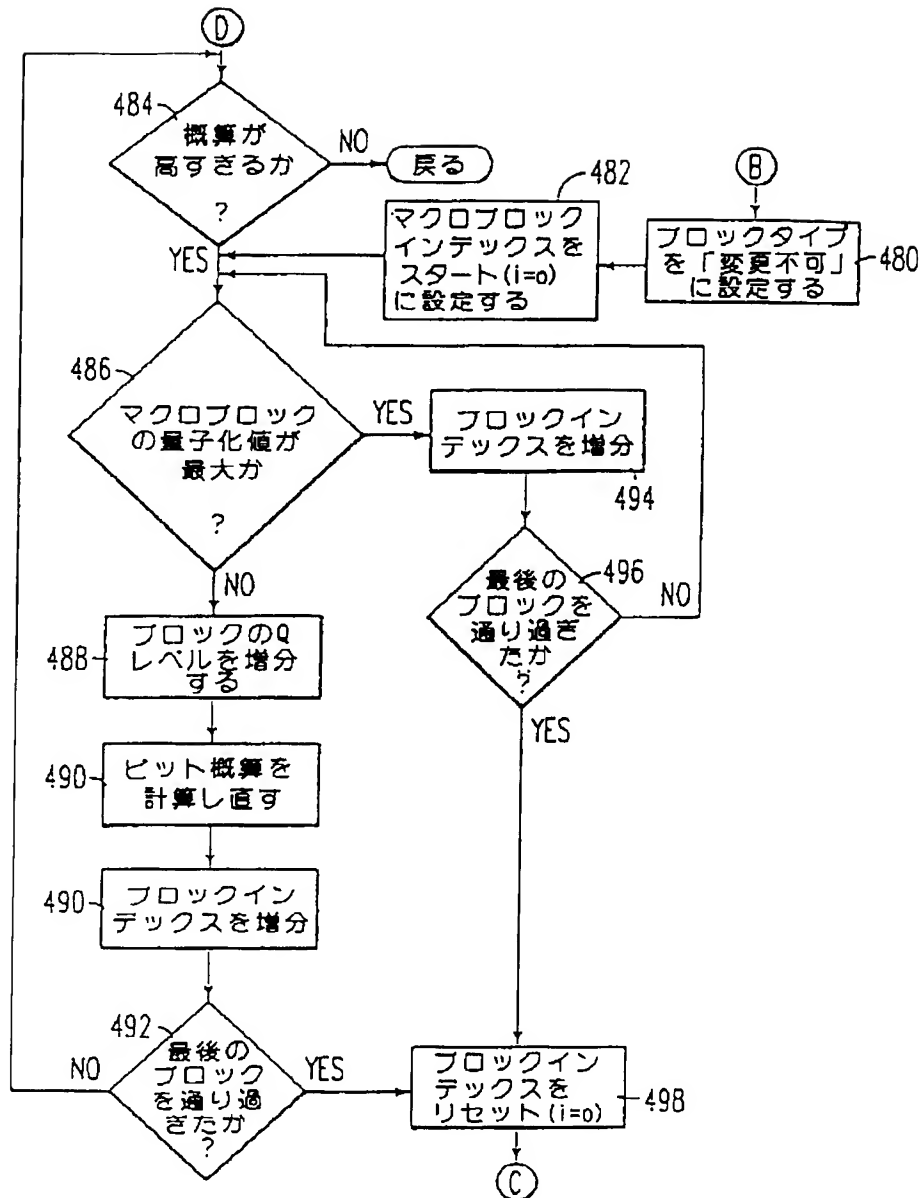
【図35】



【図20】



【図24A】



【図37】

(A)

q1	q2	q1	q2
q2	q1	q2	q1
q1	q2	q1	q2
q2	q1	q2	q1

(B)

q1	q2	q3	q4
q2	q3	q4	q1
q3	q4	q1	q2
q4	q1	q2	q3

(C)

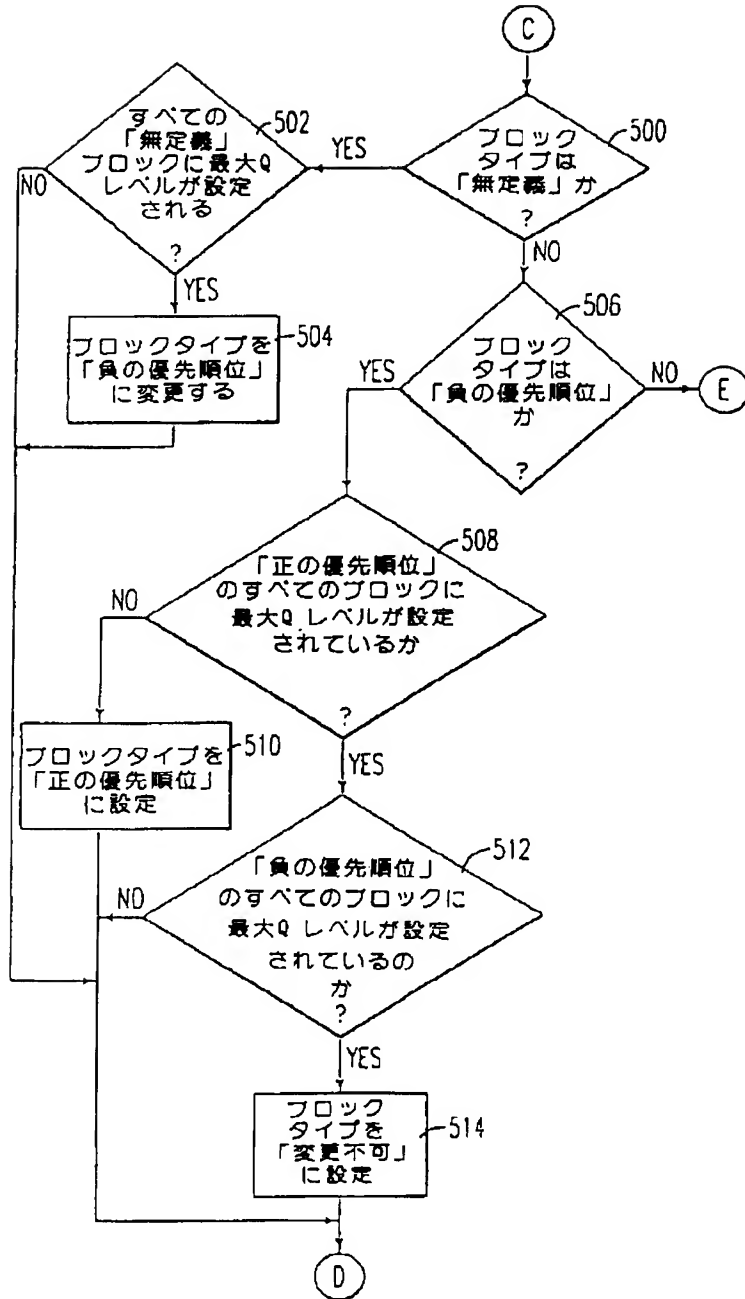
q1	q2	q1	q2
q3	q4	q3	q4
q1	q2	q1	q2
q3	q4	q3	q4

【図48】

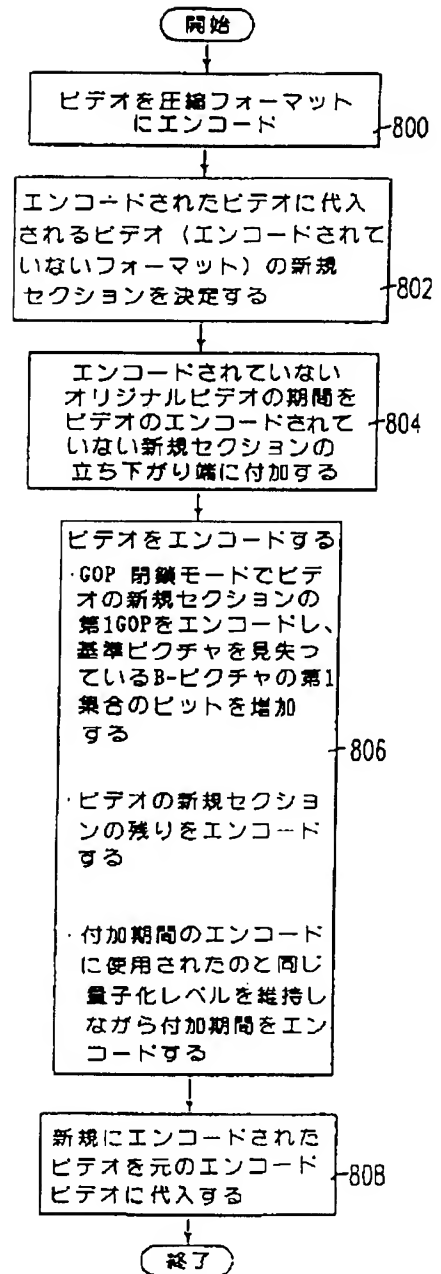
1024

オーディオバックアドレス	1026
オーディオのプレゼンテーション時刻記録(PTS)	1028
サブピクチャバックアドレス	1030
サブピクチャのプレゼンテーション時刻記録(PTS)	1032

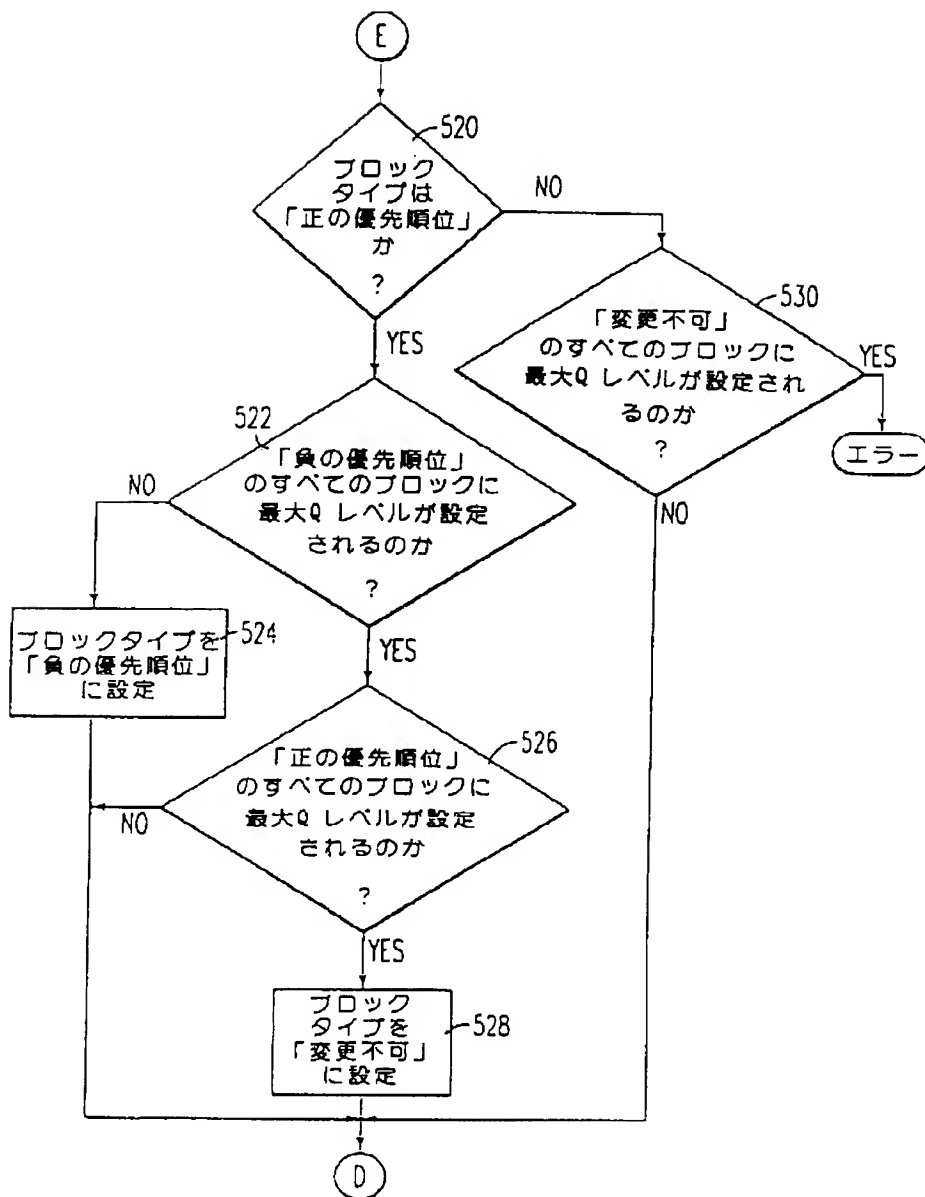
【図24B】



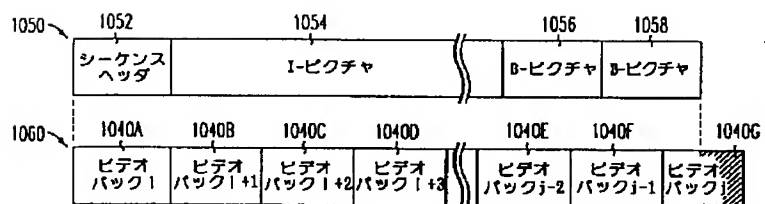
【図36】



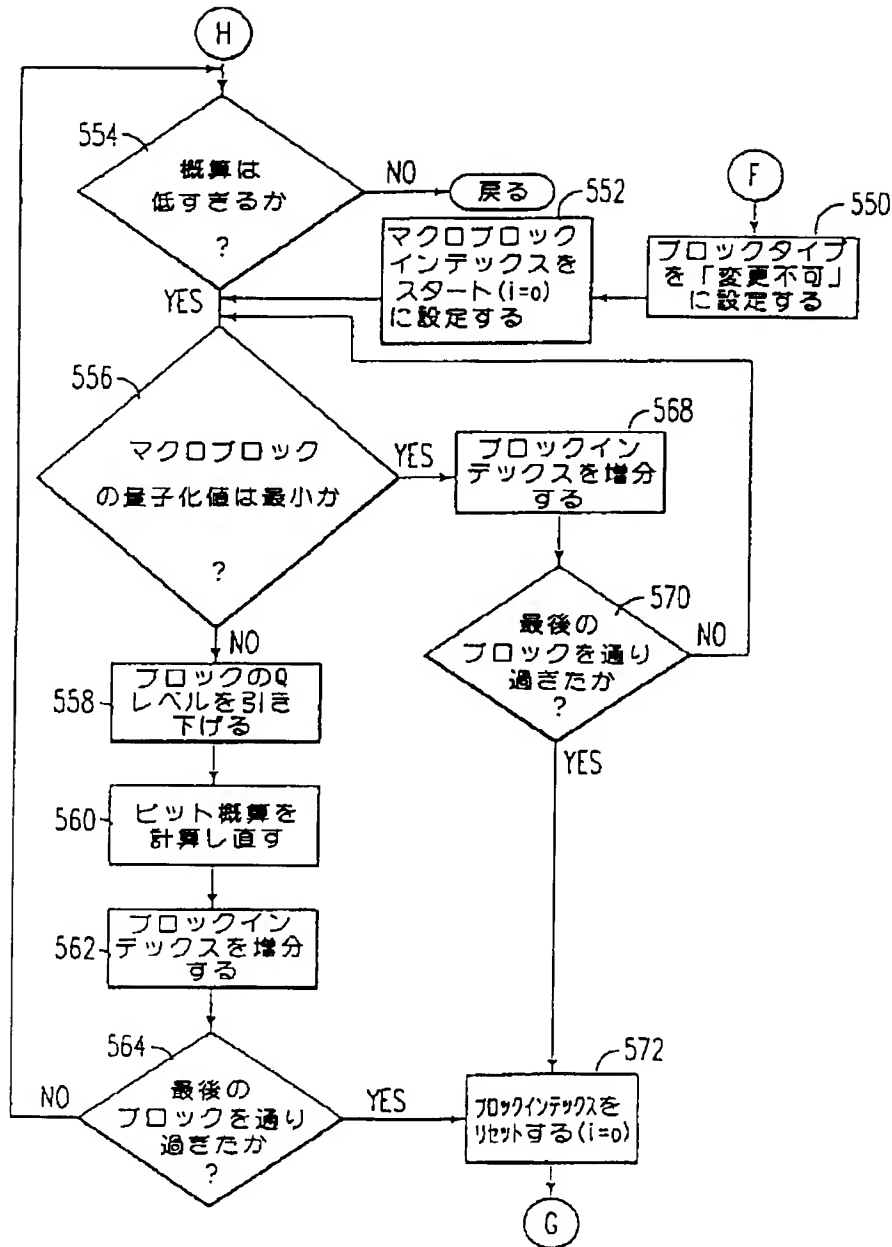
【図24C】



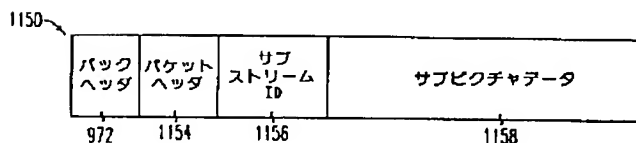
【図50】



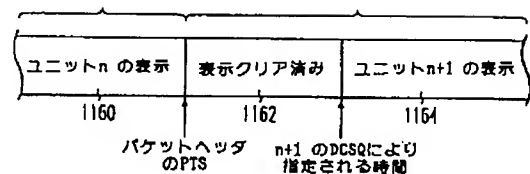
【図25A】



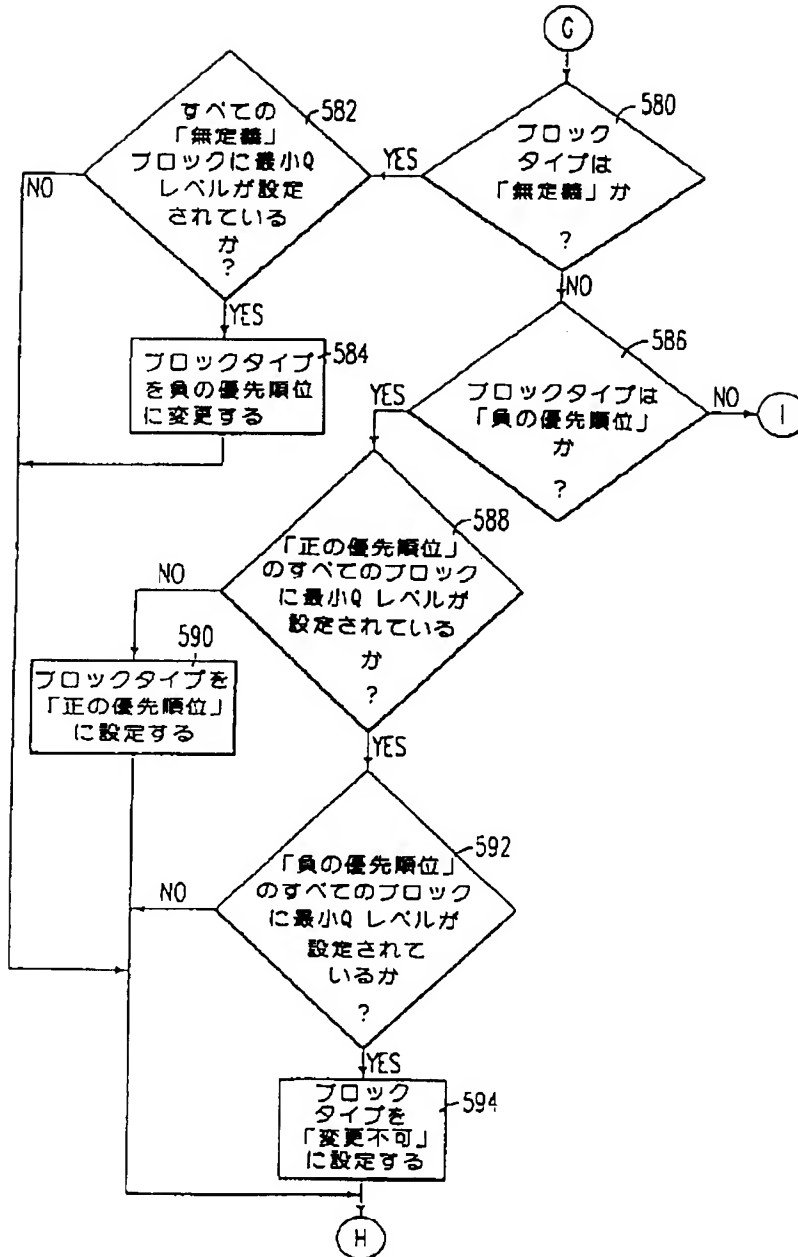
【図54】



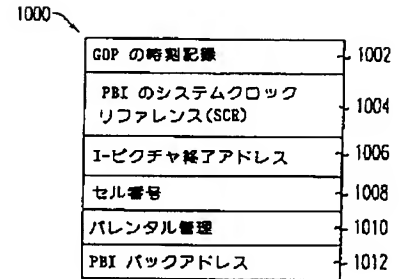
【図56】



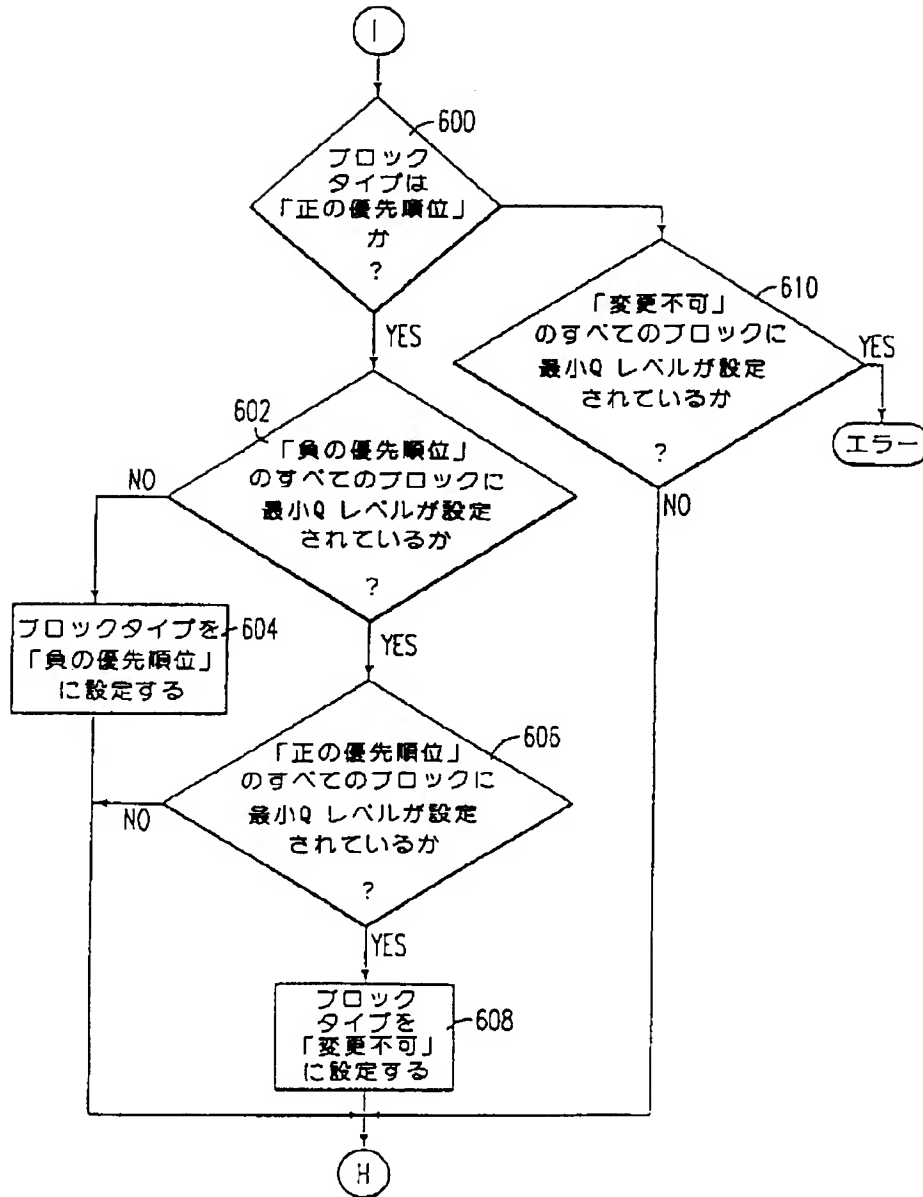
【図25B】



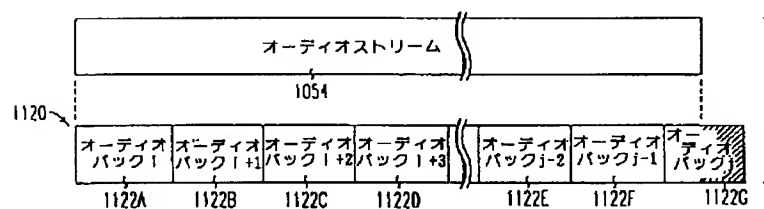
【図47】



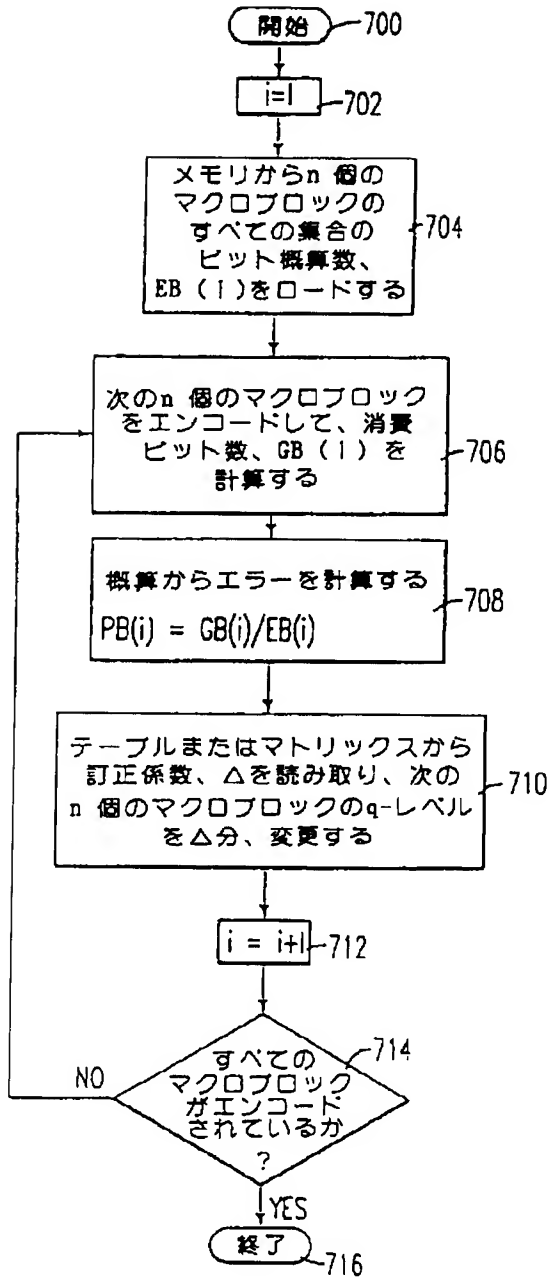
【図25C】



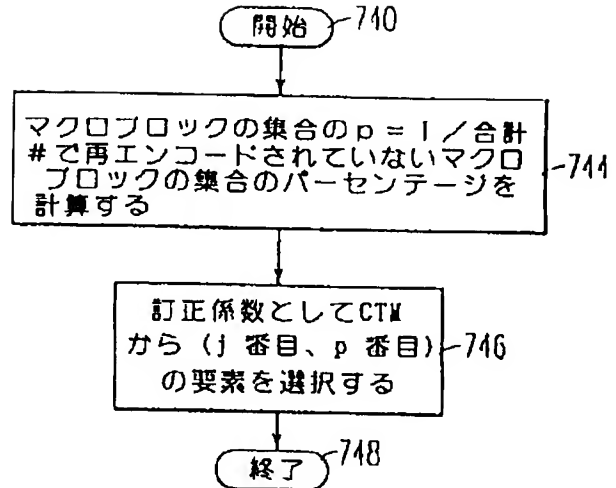
【図52】



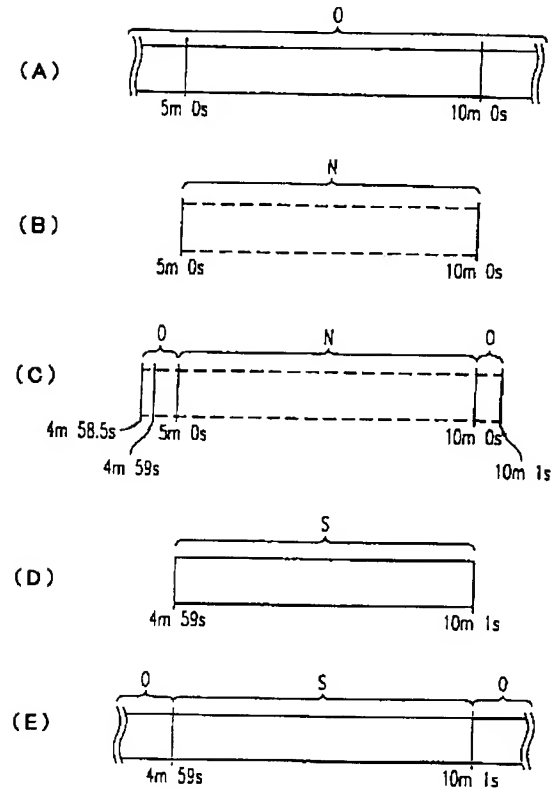
【図26】



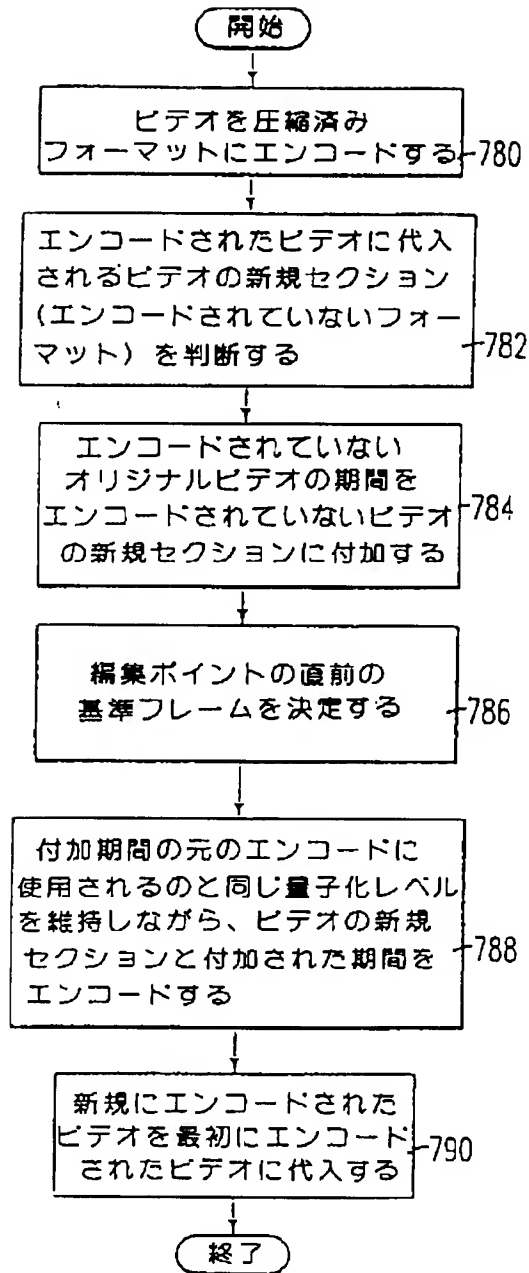
【図32】



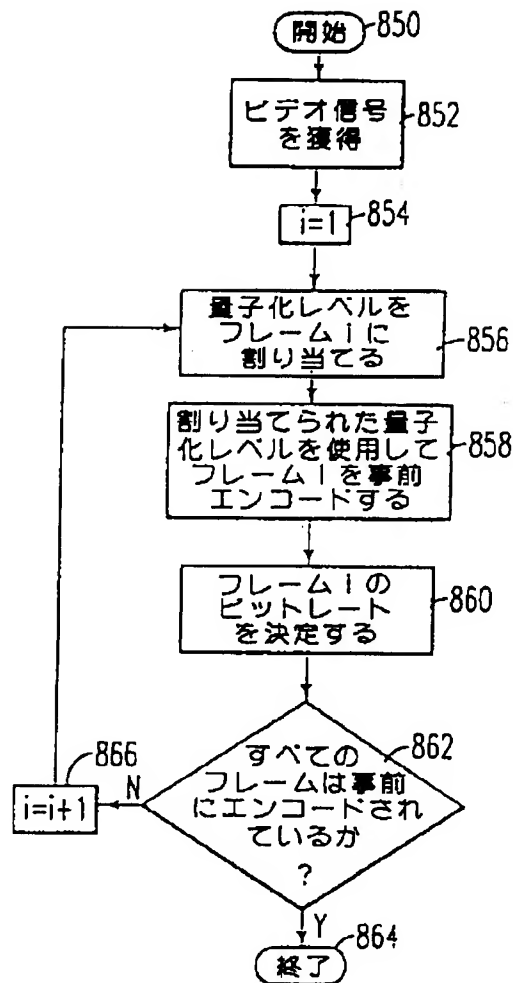
【図33】



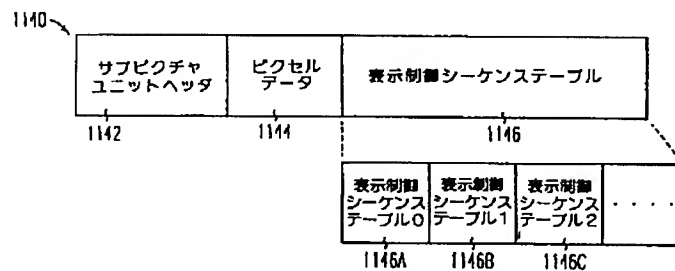
【図34】



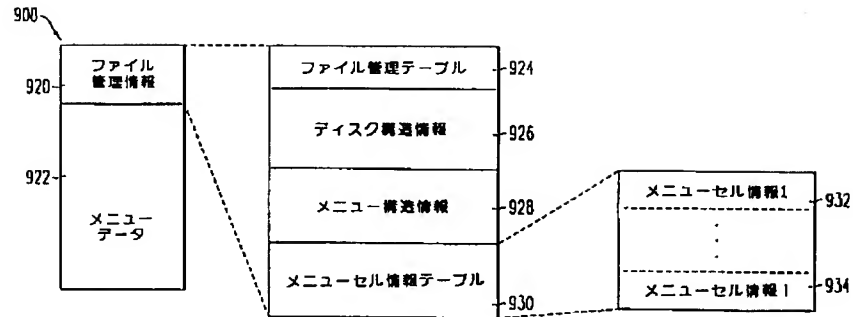
【図38】



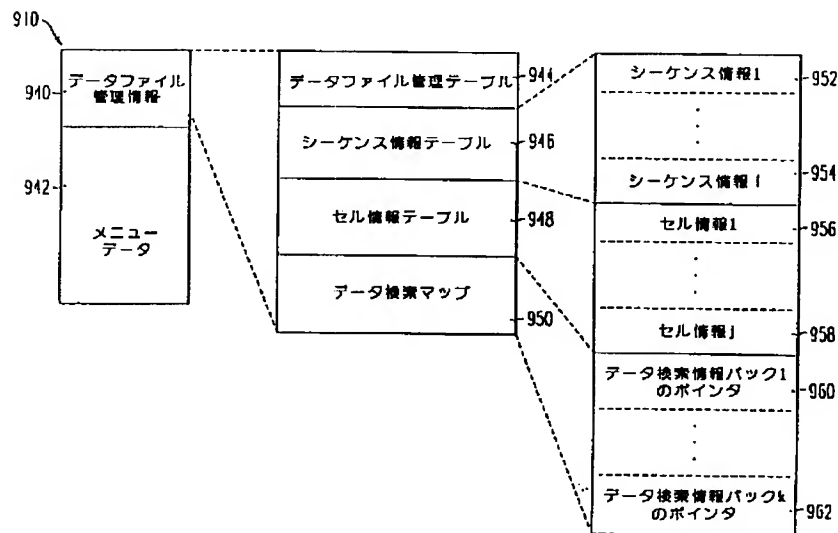
【図53】



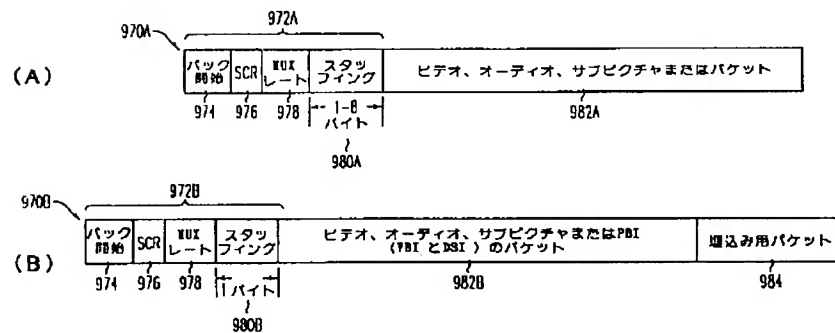
【図41】



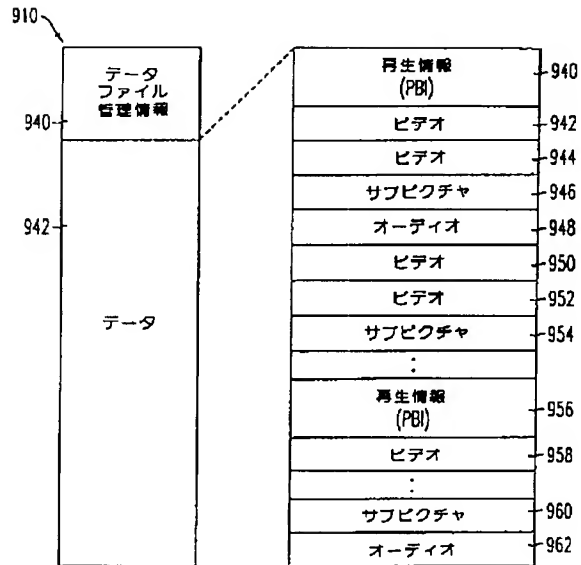
【図42】



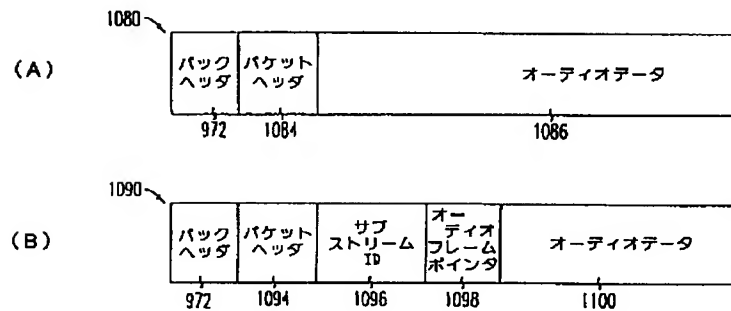
【図44】



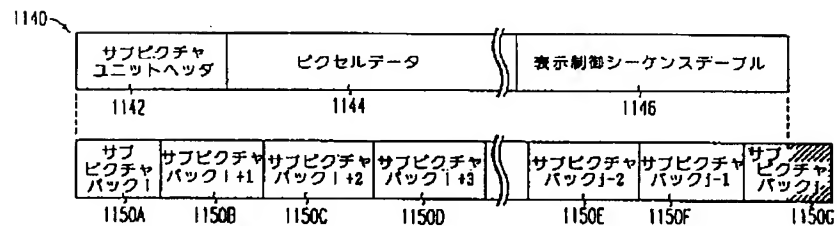
【図43】



【図51】



【図55】



フロントページの続き

(51) Int. Cl. 7

H04N 5/91
5/93

識別記号

F I

H04N 5/93

ターマコード (参考)

Z

- (72) 発明者 ジェイ・ヨーグシュウォー
アメリカ合衆国、ニュージャージー州
08648、ローレンスビル、アンソニー・レ
ーン 34
- (72) 発明者 シャウーバオ・ン
アメリカ合衆国、ニュージャージー州
08512、克蘭ベリー、ペティー・ロード
4
- (72) 発明者 伊知川 慎一
日本国神奈川県横浜市金沢区並木 3-2-
7-623
- (72) 発明者 海野 裕明
日本国千葉県市川市八幡 1-20-9
- (72) 発明者 三村 英紀
日本国神奈川県横浜市金沢区並木 2-3-
1-403
- (72) 発明者 北村 哲也
日本国東京都狛江市中和泉 1-15-12-
301
- (72) 発明者 クリストファー・ジェイ・クックソン
アメリカ合衆国、カリフォルニア州
90046、ロサンゼルス、トーレイソン・ド
ライブ 7825
- (72) 発明者 グレグ・ビー・サガード
アメリカ合衆国、カリフォルニア州
90277、レドンド・ビーチ、エメラルド・
ストリート 601
- (72) 発明者 アンドリュー・ドルージン・ローゼン
アメリカ合衆国、カリフォルニア州
91304-3654、キャノガ・パーク、フォー
ルブルック・アベニュー 7815